Income Risk, Habit Formation, and Precautionary Savings: The Case of Rural Households

Aditya R. Khanal[1,*] and Ashok K. Mishra[2]

Selected Paper prepared for presentation at the Agricultural and Applied Economics Association's 2016 AAEA Annual Meeting, Boston, MA, July 30-August 2, 2016.

[1] Assistant Professor, Department of Agricultural and Environmental Sciences, Tennessee State University, Nashville, TN;
[*]Correspondence: akhanal1@tnstate.edu.
[2] Kemper and Ethel Marley Foundation Chair and Professor, Morrison School of Agribusiness, Arizona State University, AZ.

Income Risk, Habit Formation, and Precautionary Savings: The Case of Rural Households

1. Introduction

The life-cycle hypothesis implies that individuals plan their consumption and savings over time horizon (life) and smooth their consumption in the best possible way. Under life cycle or permanent income hypothesis (Modigliani and Brumberg, 1954; Friedman, 1957) consumption decision is intertemporal allocation of the resources available during lifetime; representative consumer tries to maximize his/her utility by choosing an optimal level of resources in each period, subject to lifetime budget constraint. A portfolio of literature has empirically examined this theory (Hall, 1978; Flavin 1981; Hall and Mishkin 1982; Kazarosian 1997; Mishra, Uematsu, and Powell, 2012). One of the challenges to the life-cycle hypothesis is the prospect of risk and uncertainty. The main effect of risk and uncertainty would be to generate a demand for precautionary savings. Under liquidity constraints, consumption growth should only be sensitive to increases in income because consumers can smooth consumption by savings if they expect a decrease in future incomes. On the other hand, consumers with access to credit should follow conventional life-cycle model predictions. However, empirical evidences prone deviations from this prediction—for example, Garcia et al. (1997) and Shea (1995) for credit unconstrained consumers, found that level of consumption is affected by negative realization in income. The most plausible explanation for such anomaly is the presence of asymmetric preferences—if preferences exhibit inertia, typically in the case of habit formation, households will adjust their consumption slowly (Deaton 1992; Meghir and Weber 1996).

Only a few studies have examined consumption decision in the case of habit formation and in the presence of uncertainty[3]. Allessie and Lusardi (1997) showed that consumption not only depends on permanent income and income risk but also on past consumption, a case of habit formation. However, Guariglia and Rossi (2002) pointed out that a negative exponential utility function is not a well-representation of preferences because it does not rule out the possibility of a negative consumption. Guariglia and Rossi's (2002) presentation of the utility model is based on the generalization of Weil's (1993) model[4] but accounting for habit formation. Using this generalized model, Guariglia and Rossi (2002) estimated an Euler equation of consumption changes using a panel data from British Households. Our study follows this method to examine consumption-saving behavior. We use panel data from rural households in India to examine the consumption and saving behavior under weather and income risks and also allowing for habit formation.

Interestingly, very few studies have examined the consumption-saving responses due to variability in farm income. For example, Kochar (1999) examined consumption-saving behavior with respect to income shocks in agriculture using a longitudinal data and found evidence that the households may respond to crop income shocks by increasing their market (off-farm) hours of work. However, she pointed out that empirical results need to be confirmed with larger data set because her analysis is limited by small sample size. Paxson (1992) examined farmers' savings behavior in response to rainfall

[3] Theoretically, either quadratic utility function or negative exponential utility is assumed to have a closed form solution. Allessie and Lusardi (1997) derived a closed form solution allowing for habit formation and uncertainty using a negative exponential utility function

[4] Weil's (1993) model represents a hybrid model such that preferences are isoelastic intertemporally, but exponential with respect to the risk component. Guariglia and Rossi (2002) added habit formation component to Weil's model and derived a closed form solution in the presence of labor income uncertainty.

shocks, with an assumption that more variable rainfall results in more variable income. With rainfall shocks as proxy for income variability, findings suggested that farmers have a higher propensity to save out of transitory income. However, she mentioned that income variability from panel data would have been a better indicator.

We cannot undermine the importance of the investigation on consumption-savings in rural households because it has both micro- and macro- level implications. For example, policy instruments may differ based on saving and dissaving rates of famers in low income countries like India. How fluctuations in income lead to changes in consumption depends on the saving behavior of the farm households. If farmers are able to save and dis-save while adjusting consumptions one-to-one, then policies concerning income variability may be less relevant. Additionally, farm households behave differently under weather and income uncertainties (Paxson, 1992). The proportion of saving out of consumption under uncertainty could be a good guideline for policies aiming to support rural households for food security, income generation and poverty mitigation. In that we also cannot undermine the importance of an appropriate model and data to examine these behaviors. Our study overcomes this limitation by examining the rural households' consumption-saving behavior by separately analyzing the effect of both types of risks—farm income risk (income risk from agricultural production) and labor income risk (income risk in off-farm labor income). In rural areas with low or no irrigation facilities, with no proper storage and processing infrastructures, agricultural production highly depends on whether conditions. As variability in weather conditions has strong linkage with variability in agricultural income in rural areas, we have used weather risk as a proxy of farm income risk.

In the first part of our analysis, we estimated our models by using a consumption-saving model, an Euler equation that accounts for habit formation using a panel data from India for 2009-2013 periods. Through this model, we tested for pre-cautionary saving motives among rural households in India accounting for habit formation. In the second part of the paper, we estimated household's actual savings model under risks. We estimated two savings model under two different risks: weather risk (proxy for farm income risk) and labor income risk (proxy for off-farm income risk) by treating savings as a function of past savings, past incomes, income risk, and demographic factors.

2. Literature Review

Consumption and habit formation

Habit formation was first introduced in the context of demand analysis (Pollack 1970) and is mainly of two types—myopic and rational, based on the information about their own consumption. In myopic, consumers are not aware of the effects that their current consumption decisions will have on their future marginal rates of substitution between goods and as a consequence their behavior may be time-inconsistent. In the rational case, consumers are aware of habit forming effect of current consumption. Empirically, there are mix findings for- and against- the predictions of life-cycle models. Among the reported anomalies from prediction of life-cycle model, habit formation is one of the convincing arguments to justify the slow adjustments (Meghir and Weber, 1996). Habit formation relies on the idea that one's past consumption might have an effect on the utility obtained by current consumption.

Moreover, some anomalies in macro-level models that contrast with permanent income prediction can be resolved when allowing for habit formation. For example, Carroll et al. (2000) showed that “high growth leads to high saving”—a consistent finding with standard growth model—is obtained when accounting for habit formation. However the results were inconclusive when habit formation was not taken into consideration. Although there has been an increasing interest in habit formation in theory and evidences based on aggregate data, a very limited number of studies have used micro-level data to examine this behavior.

One of the common approaches in microeconometric studies to test for the presence of habit formation is through Euler equation. Guariglia and Rossi (2002) derived a closed form solution of the model under uncertainty while accounting for habit formation. They also tested the model using British household panel survey and found a significant evidence for habit formation. Among few other studies, Rhee (2004) and Alessie and Teppa (2010) found an evidence for habit formation using household level data in Korea and the Netherlands, respectively. On the other hand, Dynan (2000) found only a little evidence using a household level Panel Study on Income Dynamics (PSID) data from 1974 to 1987 in the United States. However a limited number of studies have examined consumption as well as precautionary saving motives taking into account the habit formation (McKenzie 2001; Guariglia and Rossi 2002). Our study aims to contribute to this limited empirical literature using panel survey data from rural India.

Consumption and savings in rural India

Following the economic reforms initiated in 1991, saving performance is a prominent policy debate in India (Athukorala and Sen, 2002). There has been a consistent increase in the national savings rate in India, in recent decades, but with yearly. Private savings comprises a greater share in the national savings; share of public savings, on the other hand, is declining since the 1980s. Economic growth and consumer buying capacity have accelerated sharply since 1990s (Landes and Burfisher, 2009). However, the national growth, savings, and consumption figures may not comply with rural farm sector as this sector has poor economic performance. Some studies have indicated that reduction in pervasive rural poverty in India is subject to a question despite an overall economic growth (Landes and Burfisher, 2009). Expenditure on food accounts for 47% of India's private consumption expenditures on goods and services and food accounts for a larger share of the total household expenditures in rural households. There are significant differences in rural and urban sector growth, consumption, and expenditures. Therefore, aggregate consumption-saving figures may not represent a real picture of rural Indian economy. Demand for precautionary saving is expected to rise with uncertainties in future income. In rural agricultural households, uncertainties mainly come from variability in agricultural and non-agricultural income. To the best of our knowledge, none of the previous studies have examined the consumption responses and precautionary saving behaviors in presence of risks in agricultural and non-agricultural incomes in rural India, especially during recent fast growing Indian economy. This study fills this gap by examining consumption responses under agricultural and non-agricultural income risks and tests for pre-cautionary saving motives. To do so, this study estimates Euler equation using a panel data from rural households. Additionally, this study accounts for habit formation in consumption.

3. Theoretical Framework

We follow life-cycle utility function maximization approach based on generalization of Weil's (1993) model while accounting for habit formation (Guariglia and Rossi, 2002). Let us assume that the household maximizes the utility function U, a constant relative risk aversion (CRRA) where preferences are characterized by constant elasticity of intertemporal substitution (equal to $\frac{1}{\alpha}$). Let us denote consumption, total assets (resources), and income at time t as c_t, a_t, and y_t, respectively. Let δ and R represent subjective discount and interest factors, respectively. The utility function can be shown as:

$$U(c_t^*, c_{t+1}^*, \dots \dots \dots) = \left\{(1-\delta)\sum_{s=0}^{\infty} \delta^s {c_{t+s}^*}^{1-\alpha}\right\}^{\frac{1}{1-\alpha}} \tag{1}$$

Equation (1) can be represented in recursive form:

$$U(c_t^*, c_{t+1}^*, \dots \dots \dots) = V\{c_t^*; U(c_{t+1}^*, c_{t+2}^*, \dots \dots .)\}$$

$$= \left\{(1-\delta){c_t^*}^{1-\alpha} + \delta[V(c_{t+1}^*, c_{t+2}^*, \dots \dots)]^{1-\alpha}\right\}^{\frac{1}{1-\alpha}} \tag{2}$$

Individual household maximizes utility function subject to the yearly budget constraint.

$$a_{t+1} = R(a_t - c_t) + y_{t+1} \tag{3}$$

$\lim_{i \to \infty} R^{-i} a_{t+i} \geq 0$ by transversality condition.

Let us now introduce uncertainty in the model by assuming a stochastic process for income which takes the following AR (1) with drift form.

$$y_{t+1} = \rho y_t + (1-\rho)\hat{y} + \varepsilon_{t+1} \tag{4}$$

y_{t+1} and y_t represent incomes in $t+1$ and t periods, respectively. $\hat{y}$ is the predicted component of income and the error term ε_{t+1} is assumed to be i.i.d. normally distributed with mean 0 and variance σ^2. Now assume that β represents the attitudes towards the risk, a constant positive coefficient of absolute risk aversion. Similar to Weil (1993), we also assume isoelastic utility preferences intertemporally but exponential in risk dimensions. In this framework, the certainty equivalent utility of a lottery yielding a random utility $U''$is U' as follows:

$$e^{-\beta U'} = E\{e^{-\beta U''}\} \tag{5}$$

where E represents expectation conditional on information available at time t, which follows: $U' = lnEe^{-\beta U''}/-\beta$.

Let us use the notation $U'\left(c_t^*, \widetilde{c_{t+1}^*}, \widetilde{c_{t+2}^*}, \ldots \ldots \ldots\right)$ as the certainty equivalent of the time *t+1* utility, conditional on the information available at time *t*. Now, the utility maximization can be shown as a recursive way which representing an aggregation of the current consumption c_t^* and the certainty equivalent of future utility.

$$U\left(c_t^*, \widetilde{c_{t+1}^*}, \widetilde{c_{t+2}^*}, \ldots \ldots \ldots\right) = V\{c_t^*;\ U'\left(\widetilde{c_{t+1}^*}, \widetilde{c_{t+2}^*}, \ldots \ldots ..\right)\} \tag{6}$$

The optimal solution of the above equation subject to budget constraint and transversality condition can be characterized by a value function in Bellman equation framework

$$W(a_t, y_t, c_{t-1}) = max_{c_{t>0}} \left\{\{(1-\delta)(c_t - \gamma c_{t-1})^{(1-\alpha}\} + \right.$$

$$\left. \delta[\frac{lnE\ exp\{-\beta W[R(a_t - c_t) + y_{t+1,}c_t]\}}{-\beta}]^{1-\alpha} \right\}^{\frac{1}{1-\alpha}} \quad (7)$$

The optimization solution of this Bellman equation leads to the following closed-form for c_t (For detail derivation, see Guarglia and Rossi, 2002).

$$c_t = \left[1 - R^{\frac{(1-\alpha)}{\alpha}} \delta^{\frac{1}{\alpha}}\right]\left(1 - \frac{\gamma}{R}\right)[a_t + \frac{1}{R-\rho}\rho y_t + \frac{1}{R-\rho}\frac{R}{R-1}\left((1-\rho)\hat{y} + \varepsilon^*\right)] +$$

$$\left(R^{\frac{(1-\alpha)}{\alpha}} \delta^{\frac{1}{\alpha}}\right)\gamma c_{t-1} (8)$$

This generalized consumption function represents three components: level of income and total resources component $(a_t, y_t, \hat{y})$; precautionary savings component (ε^*), and past consumption component, representing habit formation (the term in c_{t-1}). From the above equation, note that if $\gamma = 0$, preferences exhibit no habits while higher γ indicates the higher importance of habits in influencing optimal consumption. Also, note that under no uncertainty and no habit formation ($\gamma = 0$), we left with the usual closed form solution for consumption obtained in life-cycle/permanent income model. The precautionary component is given by:

$$\varepsilon^* = -\frac{\sigma^2}{2}\left[\frac{\beta R}{R-\rho}\right]\left[(1-\delta)^{\frac{1}{1-\alpha}}\left(1-\frac{(\delta R)^{\frac{1}{\alpha}}}{R}\right)^{\frac{\alpha}{\alpha-1}}\left(\frac{R-\gamma}{R}\right)\right] \tag{9}$$

As we can notice from equation (9), presence of habits affects optimal consumption not only through c_{t-1}, but also indirectly through making the precautionary component smaller. To derive empirical Euler equation of consumption in simple form, assume $\delta R = 1$. Now, we obtain:

$$\Delta c_{t+1} = \gamma \Delta c_t + \frac{R-1}{R}\left(1-\frac{\gamma}{R}\right)\left[\left(\frac{R}{R-\rho}\right)(\varepsilon_{t+1}-\varepsilon^*)\right] \tag{10}$$

where ε_{t+1} is the residual obtained from the income process described in equation (4) and ε^* is the precautionary savings component. Equation (9) suggests that ε^* is negative indicating that precautionary savings affect consumption changes positively, i.e., consumers face uncertainty by postponing consumption.

4. Econometric Specification

Based on the theoretical framework described in equation (10), consumption change in time *t* essentially depends on consumption changes in time *t-1* and on the precautionary component, ε^*. A special case of models under sequential exogeneity restrictions is autoregressive models. As shown in equation 4, we assume a AR(1).

On the basis of theoretical equation, we estimate the following Euler equation:

$$\Delta c_{it} = \alpha + \gamma \Delta c_{i(t-1)} + \beta_1 VAR_{it} + \beta_2 X_{it} + v_i + v_t + e_{it} \quad (11)$$

where Δ is the first difference operator. Δc_{it} and $\Delta c_{i(t-1)}$ represents household *i*'s average monthly consumption on food as a proxy for total non-durable household consumption in period *t* and *t-1*, respectively. VAR_{it} is the proxy for ε^*, income risk faced by the household *i* at time *t*. In our empirical estimation, we will separately estimate income variability in labor income (off-farm income) and income variability in agricultural income. Notice that the precautionary component of consumption is a function of σ^2 (equation 9), which represents the variance of the residuals in the income process described in equation (4). In computing VAR_{it}, we will obtain the residuals from random effects regression of the household's labor earnings (off-farm earnings) on lagged earnings, age, education, gender, regional dummies, occupational dummies, and the interaction of the education and occupational dummies with age. We then calculate the variance of these residuals in three or more years preceding and including year t. Utility function is likely to vary with demographics, family characteristics and other socio-economic variables since these may lead to shifts and variations in tastes. Therefore, we will add variables (X_{it}) on the right-hand side of the Euler equation. Finally, equation 11 shows three components of the error term: a household specific term v_i, time specific term v_t, and idiosyncratic term e_{it}. We account for the time specific effect by including year dummies in all our specifications.

We compute Euler equation using dynamic panel data models—pooled ordinary least squares (POLS), within groups estimator and generalized method of moments (GMM). We compute difference GMM and system GMM estimators designed and mostly suited for "small T, large N" (small time period) panel data analysis.

Pooled OLS method ignores time dimension and treats the data as cross-sectional by pooling across years. In the context of panel data, we usually must deal with unobserved heterogeneity. One of the common methods to deal with unobserved heterogeneity is to do the within (demeaning) transformation, the one-way fixed effect models, or by taking first differences. However, a difficulty arises due to demeaning in “small T, large N” type of panels, in particular because demeaning subtracts individual’s mean value of dependent variable from each independent variables which may create a correlation between regressors and error term (Nickell, 1981). Our estimators embody the assumptions of habit formation (process is dynamic with current realization in dependent variable (consumption) is influenced by past ones), and idiosyncratic disturbances are uncorrelated across individuals. We tested for serial correlation, used instruments, and tested for overidentifying restrictions.

Additionally, the ability of first differencing to remove unobserved heterogeneity also underlies the family of estimators that have been developed for dynamic panel data (DPD) models—difference GMM (Arellano–Bond estimators, Arellano and Bond, 1991) and system GMM estimators (Arellano and Bover, 1995 and Blundell and Bond 1998). A key feature of the Arellano–Bond and Arrellano-Bover/ Blundell-Bond estimators is that they allow internal instruments (based on lagged values of the instrumented variable(s)) as well as external instruments (other instruments in addition to lagged value). These embody the ‘first-differencing’ method that essentially removes the error term and its associated omitted-variable bias.

The Arellano–Bond estimator sets up a generalized method of moments (GMM) problem in which the model is specified as a system of equations, one per time period, where the instruments applicable to each equation differ (for instance, in later time periods, additional lagged values of the instruments are available). A potential weakness in the Arellano–Bond estimator was revealed in later work by Arellano and Bover (1995) and Blundell and Bond (1998). The lagged levels are often rather poor instruments for first differenced variables, especially if the variables are close to a random walk. Their modification of the estimator includes lagged levels as well as lagged differences. While Arellano–Bond estimator is often referred difference GMM method, Arrellano-Bover/ Blundell-Bond estimators are commonly termed as System GMM.

Difference GMM and system GMM methods are referred as the most suitable methods developed so far for dynamic panel data with unobserved heterogeneity. As these estimators are instrumental variables methods, it is particularly important to evaluate the Sargan–Hansen test results when they are applied. Another important diagnostic is the AR test for autocorrelation of the residuals. By construction, the residuals of the differenced equation should possess serial correlation, but if the assumption of serial independence in the original errors is warranted, the differenced residuals should not exhibit significant AR (2) behavior. If a significant AR (2) statistic is encountered, the second lags of endogenous variables will not be appropriate instruments for their current values.

Data and variables

To conduct empirical analysis in this study, we obtained rural household level panal data from India for the years 2009-2013 collected by International Crops Research Institute for Semi-Arid Tropic (ICRISAT) as part of Village Dynamics Study in South Asia (VDSA) program. ICRISAT micro-level data contains information on production, price, markets, climate and socio-economic aspects from representative villages across India. This study uses farm households from 18 villages in five different states, namely Andhra Pradesh (AP), Madhya Pradesh (MP), Maharashtra (MH), Gujarat (GJ) and Karnataka (KT). The villages were selected randomly to be representative of different agro-ecological zones in India. In each village, 40 sample households were selected representing households in labor/landless, small farms, medium farms, and large farm category (see Jodha et al., 1977). These initial 40 households along with some additional households in some years were surveyed and tracked over years. The households analyzed in this study are rural households that represent mainly the farm households in low-income economy. Households rely primarily on various agricultural incomes but also have, to various degrees, non-agricultural or off-farm income.

The information on ICRISAT data are presented under different modules. Consumption on food and non-food items are presented under *transaction* module. Total annual incomes, farm and non-farm incomes, livestock and land holdings, and other demographic information are arranged under *general endowment* model. Consumption variable in this study is the annual total consumption in food. Total labor income (non-farm income) variable in this study is the sum of incomes from wages received in cash or kind, and services or labor. The uncertainties in labor income (Labor income uncertainty, VAR_{it}) in the data set is the residual from a random effects regression of the household's labor earnings. To obtain this, we first estimated the household's labor income model by regressing labor earnings on lagged earnings, age of the household head, age-squared, gender, regional dummies, education dummies (primary school, high school, college and above), occupational dummies, and interactions of age and occupation with education. Then, we computed the variance of the residuals from labor income model in the two or more years preceding and including year t. For example, VAR_{i2010} (labor income uncertainty for household *i* in year 2010) is the variance of the residual from labor income model (regressing labor earnings on lagged earnings, and demographic variables mentioned above) computed based on 2009 and 2010 data.

In the second part, we estimate the relationship between actual savings with weather risk and income risk. Dependent variable in the second part is household *i*'s annual savings in year *t*. Total annual savings are computed by subtracting total annual food and non-food expenditures from total annual incomes. To compute weather risk variable, we collected weather-related information from ICRISAT *meso*-level (macro, district level) data sets. ICRISAT *meso*-level data set provides information on annual rainfall since 1966 at the district level. We computed weather risk as a coefficient of variation (CV) of rainfall based on historical district-level data. CV for respective years are computed based on the annual rainfall 40 years preceding the year. For example, CV for 2009 is based on rainfall data from 1966 to 2008; CV for 2010 is based on rainfall data from 1967 to 2009 and so on. Then district level CV variable matched with respective household in the district is merged with micro-data. Labor income risk in the second part of our estimation is the coefficient of variation in total labor income of the household. To compute CV for labor income, we collected household's total annual wage and labor income data from employment schedule (K-module) from the data window 2005-2008 for the common households. Then CV for respective years were computed from 5 years preceding and including this year. For example, CV for 2009 is computed based on labor income data from 2005 to 2009; CV for 2010 is computed based on labor income data from 2006 to 2010, and so on.

Results and Discussion

Notice that VAR_{it} is the proxy for ε^* that we have derived in equation 9 and 10. ε^* can be a representation of the labor income risk faced by the household *i* in year *t*. The precautionary component of consumption is in fact a function of σ^2, which represents the variance of the residuals in the labor income process. As described in the data section, VAR_{it} is the variance of the residual obtained from the random effects regression of labor income process. As utility function is likely vary with variables such as family size, education, gender and age of the household head which represent shift in taste, we included these factors as independent variables in Euler equation.

Table 2 shows the estimation of equation 11 for a range of estimators. Column 2 refers to OLS levels estimates. The lagged changes in consumption have a negative and significant effect on the current change. This indicates that the coefficient γ in equation 10 and 11 is negative, suggesting that the utility function exhibits habit formation or durability (Deaton 1992). The coefficient for pre-cautionary saving component VAR_{it}, on the other hand, is non-significant. Another important variable significant in OLS regression is the college education which has a significantly positive effect on consumption changes. OLS estimator however may result in biased estimates, most likely the upward biased estimates of the autoregressive coefficients, because it does not control for possibility of unobserved household-specific effects.

Column 3 of table 2 presents estimates obtained from within group estimator taking an account for the fixed effects. The estimated value of γ is still negative but smaller in magnitude than the OLS. As suggested by Nickell (1981), the fixed effect or within group estimators in the case of data with short time dimension may likely suffer from a bias. The coefficient of VAR_{it} is now positive and significant. However, estimates may be biased if there is an endogeneity of lagged differences in consumption or have measurement error.

Column 4 of table 2 presents estimates from first-differenced Generalized Method of Moments (GMM) procedure. First differencing is useful in controlling for fixed effects and time invariant component of measurement error which may likely affect both VAR_{it} and $\Delta c_{i(t-1)}$. However, time variant measurement error and endogeneity problem associated with VAR_{it} and $\Delta c_{i(t-1)}$ may still remain. Therefore, we instrumented these variables with appropriate lags (see table note, table 2). The coefficients of $\Delta c_{i(t-1)}$ (estimated value of γ) and VAR_{it} are significant. A negative and significant coefficient on $\Delta c_{i(t-1)}$ indicates an evidence for habit formation or durability in consumption. A positive and significant coefficient on VAR_{it} is an indication for precautionary savings motive among Indian households. This finding is consistent with findings from previous studies in other countries (In UK by Guariglia, 2001 and Guariglia and Rosi 2002; in China by Giles and Woo, 2007).

Table 2 also shows the test results that we conducted to evaluate whether our model is correctly specified. We used two tests: Sargan-Hansen test (*J* statistic) for over identifying restrictions and serial correlation test for second order serial correlation of residuals (*m2*). The over-identifying restriction test allows us to evaluate the validity of the instruments. If model is correctly specified, the variables in the instrument set must be uncorrelated with error term e_{it} (equation 11). The serial correlation test of second degree (*m2*) allows us to test about the legitimacy of variables dated *t-2* as instruments. *J*-

statistics of 1.42 with p-value 0.317 and no sign for second order serial correlation at 5% level suggests that our estimation does not violate the generalization of Weil's (1993) model.

Another concern while we use instrumental variables procedure is weak instrument problem. In the case of weak instruments, the estimates are likely to be biased. Particularly, Blundell and Bond (1998) indicated this issue for first-differenced GMM and developed an alternative estimator, the system GMM, which consists of combining the first-differenced equations with equations in levels. Although we do not have a prior knowledge to believe that our model is affected by weak instrument bias, we presented in column 5 of table 2, the estimates of equation 11 using system GMM estimator. We can see that the coefficients on $\Delta c_{i(t-1)}$ and VAR_{it} are very similar to those obtained from first-differenced GMM estimator. Finally, we performed a regression based formal test by regressing first differences of the potential endogenous variables $\Delta c_{i(t-1)}$ and VAR_{it} on all exogenous variables and remaining instruments using random effects model and found that the instruments have the high explanatory power.

Column 2 of table 3 presents results of the first-differences GMM estimation of Euler equation (Weil, 1993) without including habits. The serial correlation test (m2) shows the presence of second-order serial correlation of the residuals indicating that the model would be poorly specified if we do not account for the habit. Finally, we investigate whether consumption changes is excessively sensitive to income change. We investigated this by including changes in annual household income in equation 11. To control for potential endogeneity of income, we instrumented the variable with its lags (see notes on table 3). The first differenced GMM estimator of this new Euler equation are presented in column 3 of table 3. Note that the coefficient of income variable is not significant indicating that the model does not suffer from excess sensitivity of consumption to income changes. The coefficients on $\Delta c_{i(t-1)}$ and VAR_{it} are comparable to those from table 2.

In the second part of the paper, we estimated actual savings equations. We estimated two savings equations under model 1 and model 2. In model 1, we included weather variability (weather risk) as a proxy for income variability and evaluated rural household's savings response under weather risk. In model 2, we included labor income risk (variability in off-farm income) as a proxy for income variability and evaluated rural household's savings response. In both models, we control for lagged savings, lagged incomes, and demographic variables. Savings variable for household *i* in our analysis is computed as a difference between household's total incomes and total food and non-food expenditures. Weather risk and income risk variables are measured by the coefficient of variations from historical data (see data section for detail).

Table 4 presents the results from the savings equations estimated using three different methods: pooled ordinary least squares (OLS), generalized estimating equation (GEE) and dynamic generalized least squares (dynamic GLS). Though OLS does not account for the panel nature of the data, we have included this as a benchmark model. GEE and dynamic GLS are both consistent estimators for panel data with exogenous regressors. Table 4, columns 2, 3, and 4 show the results of model 1 across three different methods. Our results across three different methods consistently show that weather risk significantly influences the savings of the rural households. A coefficient estimate around 0.5 on weather risk variable indicates that a one point increase in standard deviation over the mean of the annual rainfall is associated with 50% more savings among rural households. These findings are consistent with findings from Paxson (1992) who found that farmers have higher propensity to save out of transitory income when they are subjected to higher rainfall variability. Additionally, we found that current savings have significant positive association with lagged savings and lagged incomes. This is consistent with Alessie and Lusardi (1997) who found that the current savings depend not only on income changes and income risks but also past savings. Finally, coefficient on female variable, significant at 10%, suggest that female headed households have less annual savings than male headed households.

Columns 5, 6, and 7 in table 4 present our results for savings equation under model 2. The coefficient on labor income risk are significant at 10% when we estimated savings equations using GEE method (column 6) and dynamic GLS (column 7). The result suggests that after controlling for lagged savings and incomes, labor income risk positively influences the savings of rural households. This indicates that with realization of higher labor income risk from the past and perhaps anticipating similar risk, households have tendency to save more. Coefficient estimates of 0.045 and 0.095 on labor income risk variable (table 4, column 6 and 7 under model 2) suggests that a one point increase in standard deviation from mean of labor income is associated with 4 to 9% additional savings among rural households. However, notice that the magnitude of coefficient on labor income risk (model 2) is significantly lower than that of the weather risk (model 1). This is plausible among rural Indian households because majority of them are dependent on agricultural income where weather risk plays an important role. Finally, we found that education of the household head positively influences the saving behavior under labor income risk. A coefficient estimate of 0.02 to 0.04, depending on the estimator, on education variable (table 4, columns 5, 6, and 7) suggests that an additional year of education of the household head is associated with around 2 to 4% increase in savings under labor income risk among rural Indian households.

Conclusion

We analyzed consumption and saving decisions of rural households using panel data methods and the data from farming households in India. In the first part of the paper, we tested for pre-cautionary saving behavior of the households while accounting for habit formation. We derived an Euler equation (Weil 1993; Guarglia and Rosi 2002) where consumption changes are a function of past consumption and labor income uncertainty. Using dynamic panel data methods, we found that lagged changes in consumption has the significant negative effect on current changes. This indicates that the utility function of rural households exhibits habit formation (Deaton 1992). Moreover, the labor income uncertainty has significant positive effect on consumption change which indicates the farm household’s motive for pre-cautionary saving (Guarglia and Rosi 2002). The results about habit formation and pre-cautionary savings are consistently significant across different panel model estimators. Our result also indicates that the model without accounting for habit formation would lead to biased estimates. Additionally, our test results suggest that consumption changes do not suffer from excess sensitivity to income changes in our model. The findings about habit formation indicates that the preferences are not seperable over time and the consumption studies with assumption of seperable preferences may lead to biased estimates—a finding that empirically contradicts to life cycle model.

In second part of the paper, we estimated saving equations treating the household's annual actual savings as a function of past savings, past incomes, weather and income risks and household demographics. We estimated two different models: model 1 under weather risk and model 2 under labor income risk. Our findings indicate that the rural households have tendency to save more under the realization or anticipation of both weather risk and labor income risk. However, the effect of weather risk on savings is significantly larger than that of labor income risk. Additionally, the years of education of the household head positively influence savings under labor income risk. The tendency to save higher under both weather risk and labor income risk suggests that farmers are using savings to some extent as a device to insure against risks. Note that majority of rural households in our sample are dependent on agriculture and rely highly on weather conditions. In that, higher sensitivity to weather risk is plausible. Overall our analysis suggest some insightful results. First, our findings about response to weather risk suggests that the programs supporting to stabilize income or enhance consumption for rural households should focus on agricultural diversification, weather based insurance, risk management and crop loss minimization. Second, a positive relationship between labor income risk and savings suggest that alternative income generation activities and off-farm works may enable farm households to enhance current consumption by reducing postponed consumption.

References

Alessie, R., and A. Lusardi. Consumption, Saving and Habit Formation." *Economics Letters* 55.1 (1997): 103-108.

Alessie, R., and F. Teppa. Saving and Habit Formation: Evidence from Dutch Panel Data. *Empirical Economics* 38.2 (2010): 385-407.

Arellano, M. and S. Bond. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, (1991): 277-97.

Arellano, M. and O. Bover. Another Look at the Instrumental-Variable Estimation of Error-Components Models. *Journal of Econometrics*, 68 (1995): 29-52.

Athukorala, P., and K. Sen. *Saving, Investment, and Growth in India*. Oxford University Press, 2002.

Blundell, R. and S. Bond . Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, 87, (1998): 115-43.

Carroll, Christopher D., J. Overland, and D.N. Weil. Saving and Growth with Habit Formation. *American Economic Review* (2000): 341-355.

Deaton, A. *Understanding Consumption*, Oxford University Press, Oxford, 1992.

Dynan, K. E. Habit Formation in Consumer Preferences: Evidence from Panel Data. *American Economic Review* (2000): 391-406.

Flavin M. The Adjustment of Consumption to Changing Expectations about Future Income. *Journal of Political Economy*, 89, (1981): 974-1009.

Friedman, M. *A Theory of the Consumption Function*, Princeton University Press, Princeton, NJ, 1957.

Garcia, R., Lusardi, A., and S. Ng. Excess Sensitivity and Asymmetries in Consumption: An Empirical Investigation. Journal of Money, Credit, and Banking, 29, (1997): 154-76.

Giles, J., and K. Yoo. Precautionary Behavior, Migrant Networks, and Household Consumption Decisions: An Empirical Analysis Using Household Panel Data from Rural China. *The Review of Economics and Statistics* 89.3 (2007): 534-551.

Guariglia, A., and M. Rossi. Consumption, Habit Formation, and Precautionary Saving: Evidence from the British Household Panel Survey. *Oxford Economic Papers* 54.1 (2002): 1-19.

Guariglia, A. Saving Behaviour and Earnings Uncertainty: Evidence from the British Household Panel Survey." *Journal of Population Economics*14.4 (2001): 619-634.

Hall, R. Stochastic Implications of the Life-cycle/Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy*, 86, (1978): 971-87.

Hall, R. and F. Mishkin. The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households', *Econometrica*, 50, (1982): 461-82.

Jodha, N.S., M. Asolan and J. G. Ryan. Village Study Methodology and Resource Endowments of the Selected Villages in ICRISAT's Village level studies. Economic Program Occasional Paper 16. Hyderabad: ICRISAT. 1977.

Kazarosian, M. Precautionary Savings—A Panel Study. *Review of Economics and Statistics* 79.2 (1997): 241-247.

Kochar, A. Smoothing Consumption by Smoothing Income: Hours-of-work Responses to Idiosyncratic Agricultural Shocks in Rural India. *Review of Economics and Statistics* 81.1 (1999): 50-61.

Landes, M. R., and M. E. Burfisher. Growth and Equity Effects of Agricultural Marketing Efficiency Gains in India. ERR 89, U.S. Dept. of Agriculture, Economic research Services. November 2009.

McKenzie, D. Consumption Growth in a Booming Economy: Taiwan 1976-96. *Yale Economic Growth Center Discussion Paper* 823 (2001).

Meghir, C. and G. Weber. Intertemporal Nonseparability or Borrowing Restrictions? A Disaggregate Analysis using a U.S Consumption Panel. *Econometrica*, 60, (1996): 1151-81.

Mishra, A K., H. Uematsu, and R. Powell. Precautionary wealth and income uncertainty: a household-level analysis. *Journal of Applied Economics* 15.2 (2012): 353-369.

Modigliani, F. and R. Brumberg. Utility Analysis and the Consumption Function: An Interpretation of the Cross-Section Data, In: K. Kurihara (ed.), Post-Keynesian Economics, Rutgers University Press, New Brunswick, NJ, 1954.

Nickell, S. Biases in Dynamic Models with Fixed Effects. *Econometrica,* 49, (1981): 1417-26.

Paxson, C. H. Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand. *The American Economic Review* (1992): 15-33.

Pollak, R. A. Habit Formation and Dynamic Demand Functions. *Journal of political Economy* 78.4 (1970): 745-763.

Shea, J. Myopia, Liquidity Constraints, and Aggregate Consumption: A Simple Test. *Journal of Money, Credit, and Banking*, 27, (1995): 798-805.

Weil, P. Precautionary Savings and the Permanent Income Hypothesis. *Review of Economic Studies*, 60, (1993): 367-81.

Table 1: summary statistics

Variables	Definition	Mean	Standard deviation
Consumption (c_{it})	Total annual consumption in food, in thousands	36.20	20.14
Δc_{it}	Annual food consumption, first difference	9.639	73.05
Owned plots	Total value of the plots owned by the household, in thousands	358.73	586.85
Income	Total annual household Income, in thousands	147.09	217.63
Labor income (non-farm income)	Total annual non-farm income, sum of incomes from wages received in cash or kind, and services or labor, in thousands	52.40	101.75
Labor income uncertainty[1] (VAR_{it})	Residuals from random effect regressions of the household's income generation process (for precision, we divided it by 1000)	15.47	201.38
Family Size	Number of family members in a household	5.19	2.31
Female	= 1 if household head is female, 0 else	0.08	0.27
Age	Age of the head of household, years	49.13	12.48
Education	Years of education of the head of household	4.94	4.75
Year 2009	=1 if year is 2009	0.16	0.33
Year 2010	=1 if year is 2010	0.21	0.43
Year 2011	=1 if year is 2011	0.21	0.43
Year 2012	=1 if year is 2012	0.21	0.43
Year 2013	=1 if year is 2013	0.21	0.43
	Part II		
Savings	Total annual savings (Total income less total food and non-food expenditures)	111,555.70	139,730.20
Risk related variables			
Weather risk	Coefficient of variation in annual rainfall[2]	0.61	0.18
Labor income risk	Coefficient of variation in total labor income (wages in cash and kind) in a household[3]	0.68	0.47
N		3575	

[1]VAR_{it} is the residuals from a random effects regression of the household's labor earnings. Residuals is obtained from regression of the household's labor earnings on lagged earnings, age of the household head, age-squared, gender, regional dummies, educational dummies, occupational dummies, and interactions of age and occupations with education dummies. We then calculated the variance of these residuals in the 2 or more years preceding and including year *t*.

[2]coefficient of variation (CV) of rainfall is computed based on historical district-level annual rainfall data. CV for respective years are computed based on the annual rainfall 40 years preceding the year. For example, CV for 2009 is based on rainfall data from 1966 to 2008; CV for 2010 is based on rainfall data from 1967 to 2009 and so on.

[3]coefficient of variation in labor income is computed based on household's total annual wage and labor income (cash and kind). To compute CV for labor income, we also collected household's total annual wage and labor income data from employment schedule (K-module) from the data window 2005-2008 for the common households. Then CV for respective years were computed from 5 years preceding and including this year. For example, CV for 2009 is computed based on labor income data from 2005 to 2009; CV for 2010 is computed based on labor income data from 2006 to 2010, and so on.

Table 2: Euler equation estimates

	Model 1 Pooled OLS (level form)	Model 2 Within (FE) estimation	Model 3 First-differenced GMM	Model 4 System GMM
Consumption $\Delta c_{i(t-1)}$	-0.721**	-0.232**	-0.406**	-0.391**
	(24.52)	(-7.31)	(-2.14)	(-2.05)
VAR_{it} (10^{-3})	-0.002	0.004*	0.005*	0.006**
	(-0.79)	(1.69)	(1.76)	(2.08)
Household/ household head characteristics				
Value of owned plots	0.0002	0.0067**	0.066*	0.022
	(0.17)	(4.65)	(1.69)	(0.50)
Household Size	-0.266	-1.58*	-21.70**	-1.74
	(-1.03)	(-1.67)	(-2.31)	(-0.61)
Female	-1.79	49.24**	-1395.68	324.27
	(-0.51)	(3.43)	(-0.87)	(0.76)
Age of the head	0.031	11.25**	-996.77*	6.82
	(0.58)	(13.15)	(-1.72)	(1.10)
Primary education	-0.104	NS	NS	6.83
	(-0.06)			(0.06)
High school	1.620	NS	NS	102.16
	(0.98)			(1.15)
College	7.487**	NS	NS	540.17*
	(3.07)			(1.73)
More than college	0.144	NS	NS	518.69
	(0.04)			(1.32)
m2		-2.56	-1.52	-1.37
Hansen test of over-identifying restrictions	---	--	10.43	17.42
(p-value)			0.317	0.172

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$

Note: NS: non-significant. VAR_{it} is proxy for labor income risk. Dependent variable is total food consumption in thousands. Year dummies were included in each equation. A constant term was included in model 1 and 2. Model 3: Arellano and Bond (1991) estimator obtained from two-step estimation with robust standard errors and statistics. Model 4: Two-step system GMM (Blundell and Bond, 1998) estimation with robust standard errors. Instruments: Instruments in model (3): $c_{i(t-2)}$, $VAR_{i(t-2)}$, and X_{it}, household characteristics; in model (4): $c_{i(t-2)}$, $VAR_{i(t-2)}$, and X_{it} for the differenced equation; $\Delta c_{i(t-2)}$, $\Delta VAR_{i(t-2)}$, and X_{it} for the level equation. *m2* is the test for second-order serial correlation in the first differenced residuals under the null of no serial correlation.

Table 3: Euler equation estimates, alternative estimation

Dependent variable: Δc_{it}	First Differenced GMM Model 1	First Differenced GMM Model 2
$\Delta c_{i(t-1)}$		-0.236**
		(-2.01)
VAR_{it} (10^{-3})	0.003	0.008*
	(0.92)	(1.81)
Δy_{it}		-0.030
		(-1.50)
Household/ household head characteristics		
Value of owned plots	-0.017**	0.044
	(-1.99)	(0.74)
Household size	-32.36	-56.02
	(-0.51)	(-1.41)
Female	-269.81	2254.65
	(-0.09)	(0.73)
Age of the HH head	-296.79	-117.09
	(-0.32)	(-0.21)
m2	2.39	1.09
Sargan/ Hansen J-statistics	5.37	4.69
p-value	0.497	0.891

Note: Both models control for education dummies and year dummies in each regression. Instruments in model (1): $V_{(t-2)}$, and X_{it}; instruments in model (2): $VAR_{i(t-2)}$, $c_{i(t-2)}$, $y_{i(t-1)}$, $\Delta y_{i(t-1)}$ and X_{it}.

Table 4: Savings equations

Dependent variable: **Savings [S_{it}], in log**	**Savings under weather risk** (Model 1)			**Savings under labor income risk**[¥] (Model 2)		
	Pooled OLS	Population averaged GEE	Dynamic GLS	Pooled OLS	Population averaged GEE	Dynamic GLS
[a]Lagged savings [$S_{i(t-1)}$]	0.257**	0.322**	0.0443	0.745**	0.826**	-0.316
	(2.27)	(2.75)	(0.33)	(2.32)	(2.69)	(-1.01)
[a]Lagged income [$y_{i(t-1)}$]	0.749**	0.687**	0.891**	0.303	0.220	0.568
	(4.48)	(3.90)	(4.29)	(0.77)	(0.58)	(1.35)
[a]Value of owned plots	-0.0301	-0.0236	-0.0270	0.0777	0.0837	0.383
	(-0.54)	(-0.42)	(-0.41)	(0.86)	(1.13)	(1.14)
[a]Weather risk	0.492**	0.486**	0.567**			
	(2.72)	(3.16)	(2.73)			
[a]Labor income risk				0.0515	0.0449*	0.0955*
				(0.65)	(1.64)	(1.73)
Female	-0.229*	-0.218*	-0.265*	0.0375	0.0431	-0.00462
	(-1.67)	(-1.74)	(-1.64)	(0.18)	(0.23)	(-0.02)
Household size	0.0164	0.0150	0.0199*	-0.0018	-0.0117	0.302*
	(0.97)	(1.00)	(1.81)	(-0.21)	(-0.39)	(1.72)
Service as main occupation	-0.139	-0.135	-0.110	-0.163	-0.147	-0.147
	(-1.23)	(-1.20)	(-0.81)	(-0.78)	(-0.75)	(-0.83)
Yrs. of Education, HH head	-0.0102	-0.00937	-0.0135	0.039*	0.039**	0.0240*
	(-1.16)	(-1.12)	(-1.33)	(1.64)	(2.12)	(1.67)
Constant	-0.0851	-0.119	0.648	-0.525	-1.209	5.507
	(-0.08)	(-0.11)	(0.47)	(-0.34)	(-0.75)	(1.28)

All models include set of year dummies in each equations; *t* statistics in parentheses; * $p < 0.10$, ** $p < 0.05$
[a]indicates variables are in log form. [¥]coefficient of variation (CV) in labor income also utilizes data from 2005-2008, thus sample in this model includes common households from that period, which brings our sample size to 1795.