Measurement Error and Its Impact on Estimates of Income and Consumption Dynamics*

Nayoung Lee†

September 2010

Abstract

This paper examines whether reported income and consumption generate biases for studies on income and consumption dynamics, using a linear measurement error model. This study finds that substantial classical measurement error exists in reported data, leading to a bias towards zero in the estimates of income and consumption persistence. Non-classical measurement error and unobserved heterogeneity offset the effect of classical measurement error. The variance of measurement error is larger in the model for income dynamics than that for consumption dynamics. This result suggests that measurement error is more prevalent and varies more across households in income than in consumption.

JEL Classifications: C81, I32, O15
Keywords: measurement error, income dynamics, consumption dynamics

---

* **Acknowledgments:** I would like to thank John Strauss for his excellent guidance. I also acknowledge the valuable comments of John Gibson, John Ham, Cheng Hsiao, David McKenzie, Roger Moon, Jeffrey Nugent, Geert Ridder, and participants at the North American Summer Meeting of the Econometric Society in Pittsburgh, the Far Eastern Meeting of the Econometric Society in Singapore, Northeastern Universities Development Consortium Conference in Boston, Southern Economic Association Annual Meeting in Washington D.C, Pacific Conference for Development Economics in San Francisco, and seminars at the University of Southern California and Chinese University of Hong Kong. All remaining errors are mine.

† **Corresponding Address:** Department of Economics, Chinese University of Hong Kong; ELB 9/F, Shatin, N.T. Hong Kong; Email: nayoung.lee@cuhk.edu.hk; Tel: (852) 2609-8004; Fax: (852) 2603-5805.

1. Introduction

Understanding poverty dynamics and economic mobility has always been of great concern for policies related to poverty reduction and inequality. Panel data surveys have accelerated the development of such studies in the last two decades. However, both reported income and consumption, which are the main variables in these studies, often have substantial measurement error. This measurement error may bias the estimated degree of income and consumption mobility and lead to inappropriate policy conclusions. This study explores the role of measurement error in the study of income and consumption dynamics when panel data are used.

The study uses data from the Korean Labor and Income Panel Study (KLIPS) over the period 1999 to 2006 and investigates a linear dynamic panel data model with a fixed effect. This paper presents the direction and magnitudes of the biases generated by potential classical and non-classical measurement error in the model which shows the degree of economic mobility or persistence. This study investigates both income and consumption dynamics because this makes it possible to compare the magnitudes of measurement error in each variable and the degree of true income and consumption persistence.

While evidence of measurement error in reported income or consumption from validation studies is not perfectly applicable to survey data for all countries, it does establish a benchmark.[1] According to the evidence, some part of recall errors are likely to be correlated with explanatory variables such as age, sex, level of education, and job type, while the rest of errors are independent of these variables (See review by Bound *et al.*, 2001). A large number of self-employed households, for example, often confuse

[1]Some studies make comparisons between administrative records and survey data, mainly for US income data (e.g. Bound and Krueger, 1991), while other studies attempt to estimate the degree of measurement error with experimental settings, mainly for consumption data (e.g. Browning *et al.*, 2003, Ahmed *et al.*, 2006).

personal and business income and expenses with lack of written records and generate recall errors in income (Deaton, 1997, Daniel, 2001).[2] Measurement error in income may also be correlated with true income, mostly due to tax reasons (Morgenstern, 1963, Deaton, 1997). Studies to date show that the popular assumption of only classical measurement error is not suitable for income or consumption.

Instead of restricting the type of measurement error to be only classical, this paper uses a linear measurement error model (Bollinger and Chandra, 2005, Kim and Solon, 2005). This is based on the empirical evidence (Bound and Krueger, 1991, Pischke, 1995) that those who earn higher than average tend to report their earning less, while those who earn lower than average tend to report higher (i.e. so called mean-reverting measurement error). When measurement error exists, the standard Arellano and Bond (1991) estimator for dynamic panel models with fixed effects is not consistent. Holtz-Eakin *et al.* (1988) suggest an estimator that is valid when measurement error is classical. This paper shows that the Holtz-Eakin, Newey and Rosen estimator is still valid even when we use a linear measurement error model allowing for non-classical measurement error. This study also suggests various specification tests that include testing for existence of random measurement error in income and consumption data.

My main finding is that there is substantial random or classical measurement error in reported income and consumption, leading to a bias towards zero in the estimates of income and consumption persistence. Specifically, this bias is about 65 percent and 54 percent of the consistent estimate of the true measure of persistence in the study of income and consumption dynamics respectively. Non-classical measurement error and unobserved heterogeneity are also found to be important and to lead to upward biases in the estimated coefficients, offsetting the effect of random measurement error.

[2] This type of recall error is more common for rural areas in low income countries where most people are agricultural producers as well as consumers. However, Smith (1997) also finds that low-skilled workers tend to underreport their irregular earning like overtime, tips and commissions for direct questions about annual earnings.

This study also identifies the variance of random measurement error. The standard deviation of the measurement error is estimated to be as large as that of the equation error for both income and consumption dynamics, suggesting that random measurement error is substantial. Interestingly, the standard deviation of the measurement error is much larger in the model for the income dynamics (.97) than that for consumption dynamics (.18). This result suggests that random measurement error is more prevalent and varies more across households in income than in consumption. This study also suggests that the standard deviation of the equation error is substantially larger in the model for the income dynamics than that for consumption dynamics. This implies that Korean households smooth their consumption relative to their income in the face of shocks.

The remainder of the paper is organized as follows: Section II reviews literature on the study of income and consumption dynamics; Section III presents the empirical model and strategies; Section IV describes the data; Section V focuses on findings; and Section VI concludes the paper.

2. Previous Studies

Numerous studies investigate the degree of income mobility with dynamic aspects, although relatively few studies have corrected for measurement error bias. Individual earnings mobility rather than household income mobility is typically estimated in the US literature, and studies heavily rely on the autocorrelated individual component model (or variance component model) advocated by Lillard and Willis (1978). The model combines a typical earnings function and an error structure allowing for an individual random effect and a first order autocorrelation of a transitory component, but no lagged dependent variable. The effect of the permanent component on earnings inequality can be identified separately from that of the transitory shock. Other studies assume unobserved heterogeneity to be a time-invariant fixed effect and typically take first differences, so that the permanent component on earnings inequality is no longer identifiable. MaCurdy (1982) uses this approach and tries to better fit the model using time series processes and taking first differences. He chooses an ARMA(1,2) specification as his favorite for the error structure of log earnings in levels.[3] However, few studies address the measurement error issue, and those that do often require administrative data to identify the effect of measurement error in reported earnings on earnings mobility (Pischke, 1995, Gottschalk and Huynh, 2006, Dragoset and Fields, 2006).[4]

Some studies allow for a more general dynamic relationship between current and previous earnings rather than confine the source of dynamics to be a serially correlated income shock or from unobserved time-varying individual heterogeneity (Holtz-Eakin *et al.*, 1988). Those studies are more pervasive in the development literature. Studies for developing countries usually estimate economic mobility with a unit of household income (i.e. per capita household income) because individual earning mobility hardly tells much about economic well-being for the poor or poverty dynamics (i.e. in and out of poverty).[5] In particular, some recent works have emphasized a nonlinear relationship between current and lagged income to identify potential poverty traps. Lokshin and Ravallion (2004) find nonlinearity in the dynamics but no evidence for the existence of a poverty trap modelling income as a polynomial of lagged income for Hungary and

[3] Abowd and Card (1989) confirm this specification, and Moffit and Gottschalk (1995) also find that the impact of temporary income shock decays within three years for 1969-1987. In addition, Meghir and Pistaferri (2004) estimate the autocorrelated individual component model allowing for educational- and time-specific differences in the stochastic process for earnings as well as for measurement error using an Autoregressive Conditional Heteroskedasticity (ARCH) specification. They find that earnings variances are heterogeneous across individuals.

[4] Pischke (1995) uses the Panel Study of Income Dynamics Validation Study (PSIDVS). In addition, Gottschalk and Huynh (2006) and Dragoset and Fields (2006) examine the effect of measurement error on several measures of earning mobility, while not relying on the autocorrelated individual component model, but their studies also require the use of the tax records in the Detailed Earnings Record (DER).

[5] See also Baulch and Hoddinott (2000) for their review of the studies of economic mobility and poverty dynamics in the development literature.

Russia. In their study, however, measurement error in income is not allowed.

Administrative data to deal with measurement error are unavailable in most developing countries, and surprisingly few studies have identified and corrected for measurement error bias by employing econometric approaches. However, Newhouse (2005) estimates the elasticity of a household's 1997 income with respect to its 1993 income in Indonesia addressing classical measurement error in income and unobserved household heterogeneity using rainfall, asset or consumption data as instruments. However, his study cannot identify the effect of classical measurement error from unobserved household heterogeneity. Antman and Mckenzie (2007b) explore the nonlinear relationship between current and lagged income also addressing unobserved heterogeneity and measurement error through a pseudo-panel approach. Their pseudo-panel approach addresses bias from measurement error under the assumption that a law of large numbers applies within a cohort, so that mean of measurement error across cohorts converges to zero in probability as the number of individuals within a cohort increases. Therefore, they have consistent estimates from measurement error but cannot identify the effect of measurement error.

Per capita household consumption is also a popular unit to estimate economic mobility or welfare especially in the development literature (e.g. Glewwe and Hall, 1998, Jalan and Ravallion, 2002, Glewwe, 2005), but the study of consumption dynamics dealing with measurement error has not yet much developed in this literature. Glewwe (2005) and Gibson and Glewwe (2005) only obtain a correlation coefficient between current and lagged consumption addressing classical measurement error in consumption using non-holiday food purchases and the body mass index (BMI) as instrumental variables.

3. Methods

3.1. Empirical Model

The basic model of income or consumption dynamics regresses either current per capita household income or consumption in natural logarithms on its lagged value, controlling for other household demographic variables and unobserved heterogeneity. The models are:

$$\ln Y_{it}^{*} = \gamma_y \ln Y_{it-1}^{*} + \beta_y' X_{it} + D_t + \alpha_i^y + \varepsilon_{it}^y \tag{1}$$

and

$$\ln C_{it}^{*} = \gamma_c \ln C_{it-1}^{*} + \beta_c' X_{it} + D_t + \alpha_i^c + \varepsilon_{it}^c, \tag{2}$$

where Y_{it}^{*} is the true per capita income and C_{it}^{*} is the true per capita consumption of household i in time period t, X_{it} is a vector of household i's demographic variables in time period t, D_t captures time-specific effects, α_i^k (k indicates income or consumption respectively) is time-invariant unobserved heterogeneity of household i, and ε_{it}^k is an error term.

However, one does not observe the true measure Y_{it}^{*} or C_{it}^{*} but rather observes Y_{it} or C_{it}. In the linear measurement error model (Bollinger and Chandra, 2005, Kim and Solon, 2005), the observed data, Y_{it} and C_{it}, and their true values, Y_{it}^{*} and C_{it}^{*}, have a linear relationship:

$$\ln Y_{it} = \lambda_y \ln Y_{it}^{*} + e_i^y + v_{it}^y, \tag{3}$$

and

$$\ln C_{it} = \lambda_c \ln C_{it}^{*} + e_i^c + v_{it}^c, \tag{4}$$

where e_i^k and v_{it}^k ($k = y$ or c) are individual-specific and random (or classical) measurement error respectively.[6] In particular, if $0 < \lambda_k < 1$, then measurement error in income or consumption is mean-reverting, and this case is consistent with the empirical findings in the studies based on validation data. On the other hand, $\lambda_k = 1$ ($k = y$ or c) implies that time-varying measurement error is purely random and that total measurement error can be decomposed into non-classical time-invariant and only classical time-varying measurement error.

In this setting, total measurement error defined as the difference between the true measure and observed data is

$$\eta_{it}^y = (\lambda_y - 1) \ln Y_{it}^* + e_i^y + v_{it}^y \tag{5}$$

or

$$\eta_{it}^c = (\lambda_c - 1) \ln C_{it}^* + e_i^c + v_{it}^c. \tag{6}$$

Here, time-invariant individual-specific measurement error e_i^k is free of assumptions (that is, e_i^k can be correlated with any other variables of the model). The non-classical part of time-varying measurement error is restricted to be correlated with only true income or consumption.[7] Nevertheless, the assumption is more plausible than assumptions which completely ignore any possibility of the existence of non-classical measurement error in time-varying measurement error. Most non-classical recall errors correlated with household characteristics are more likely to be time-invariant in this study, because it covers a relatively short time period. It is unlikely that the error would be time-varying because household characteristics are not likely to change much in such a short time.

[6] The random measurement error will be further restricted to be white noise measurement error, but this is only for the error decomposition introduced later, but this paper uses the two terms quite interchangeably.

[7] In general definitions, classical measurement error is not correlated with true income (or consumption) and other explanatory variables. Otherwise, measurement error is non-classical.

The models with measurement error, after substituting equation (3) into equation (1) and equation (4) into equation (2), are

$$\ln Y_{it} = \gamma_y \ln Y_{it-1} + \lambda_y \beta_y' X_{it} + \lambda_y D_t + \lambda_y \alpha_i^y + \lambda_y \varepsilon_{it}^y + (1-\gamma_y) e_i^y + v_{it}^y - \gamma_y v_{it-1}^y \quad (7)$$

and

$$\ln C_{it} = \gamma_c \ln C_{it-1} + \lambda_c \beta_c' X_{it} + \lambda_c D_t + \lambda_c \alpha_i^c + \lambda_c \varepsilon_{it}^c + (1-\gamma_c) e_i^c + v_{it}^c - \gamma_c v_{it-1}^c. \quad (8)$$

This paper hereafter refers to ε_{it}^k as equation error to distinguish the total composite error, $\lambda_k \alpha_i^k + \lambda_k \varepsilon_{it}^k + (1-\gamma_k) e_i^k + v_{it}^k - \gamma_k v_{it-1}^k$ for k=y or c.

The model with measurement error produces biased estimates of γ_y and γ_c, though the degree of bias depends on the assumptions of the measurement error (i.e. e and v). As is well known, the direction of the bias depends on whether the measurement error is classical or non-classical in nature. The OLS estimate of γ (either γ_y or γ_c) will be biased towards zero if only classical measurement error (i.e. v but not e) is assumed (Griliches and Hausman, 1986). However, as validation studies suggest, measurement error in reported income (or consumption) is likely to be correlated with true income (or consumption). It is thus not likely that classical measurement error is the only type of measurement error present in the estimations. In this case, the direction of the bias and the contribution of measurement errors in this level model are not theoretically determined because those of non-classical measurement error are unascertained.

In addition, the unobserved individual heterogeneity component α_i^y or α_i^c can be another source of the endogeneity. Notice that due to the nature of the dynamics in models (7) and (8), the regressors $\ln Y_{it-1}$ and $\ln C_{it-1}$ are correlated with α_i^y and α_i^c, respectively. This means that the pooled OLS estimates could be biased even without the measurement error.

Taking first differences to remove both time-invariant unobserved heterogeneity and individual-specific measurement error, we have

$$\Delta \ln Y_{it} = \gamma_y \Delta \ln Y_{it-1} + \lambda_y \beta_y' \Delta X_{it} + \lambda_y \Delta D_t + \lambda_y \Delta \varepsilon_{it}^y + \Delta v_{it}^y - \gamma_y \Delta v_{it-1}^y \qquad (9)$$

and

$$\Delta \ln C_{it} = \gamma_c \Delta \ln C_{it-1} + \lambda_y \beta_c' \Delta X_{it} + \lambda_y \Delta D_t + \lambda_y \Delta \varepsilon_{it}^c + \Delta v_{it}^c - \gamma_c \Delta v_{it-1}^c. \qquad (10)$$

This study estimates the persistence of income and consumption, γ_y and γ_c, in models (9) and (10) with possible sets of instruments, discussing how to test for the existence of the random measurement error v_{it}^k and for other specifications of the models.

3.2. Estimation of γ

As already noted, my study estimates a first-differenced model to eliminate the effect of time-invariant unobserved heterogeneity and individual-specific measurement error. Nonetheless, endogeneity remains due to classical measurement error and time-varying omitted variables as well as an AR(1) specification that includes a lagged dependent variable as an explanatory variable. This study employs two-step generalized method of moments (GMM) estimation, which uses instrumental variables (IV) to resolve this endogeneity problem for both investigations of income and consumption dynamics.

One set of internal instruments could be two-period and earlier lagged dependent variables (i.e. the Arellano and Bond estimator), but the two-period lagged dependent variable is invalid if random measurement error is assumed. However, the same instrument set but excluding the two-period lagged dependent variable can be still used in this case. In addition, this study uses external IVs that serve various roles.

Internal IV

First, lagged dependent variables are used as instruments. The variables of two or more lagged periods are well-known instruments in the first-differenced dynamic panel model. The typical model without measurement error assumes the sequential exogeneity of lagged dependent variables and the strict exogeneity of other explanatory variables conditional on α_i^k ($k = y$ or c). That is,

$$E[\varepsilon_{it}^y | \ln Y_{it-1,} \ln Y_{it-2,}..., \ln Y_{i0,} X_i, \alpha_i^y, e_i^y] = 0 \tag{11}$$

and

$$E[\varepsilon_{it}^c | \ln C_{it-1,} \ln C_{it-2,}..., \ln C_{i0,} X_i, \alpha_i^c, e_i^c] = 0, \tag{12}$$

where $X_i = [X_{i1}, ..., X_{iT}]$. The valid instruments for $\Delta \ln Y_{it-1} = (\ln Y_{it-1} - \ln Y_{it-2})$ or $\Delta \ln C_{it-1} = (\ln C_{it-1} - \ln C_{it-2})$ are the lagged levels $\ln Y_{it-2}, \ln Y_{it-3},....,\ \ln Y_{i1}$ or $\ln C_{it-2}, \ln C_{it-3},....,\ \ln C_{i1}$ because $E(\ln Y_{it-s} \cdot \Delta\varepsilon_{it}^y) = 0$ and $E(\ln C_{it-s} \cdot \Delta\varepsilon_{it}^c) = 0$ are assumed for $s = 2, 3, ...t-1$.

However, model (9) and (10) with random measurement error is distinguished from the typical dynamic panel model. That is, model (9) or (10) includes the lagged random measurement error differenced term, $\Delta v_{it-1}^k = v_{it-1}^k - v_{it-2}^k$ for k=y or c, in the total composite error differenced term. Since $E(\ln Y_{it-2} \cdot \Delta v_{it-1}^y) \neq 0$ and $E(\ln C_{it-2} \cdot \Delta v_{it-1}^c) \neq 0$, the lagged level $\ln Y_{it-2}$ and $\ln C_{it-2}$ are not valid instruments. The set of lagged levels $\ln Y_{it-3},....,\ \ln Y_{i1}$ or $\ln C_{it-3},....,\ \ln C_{i1}$ is a valid instrument set to correct for random measurement error. The difference of coefficients between models using the first and the second set of instruments indicate the direction and contribution of any bias by random measurement error because the first set of instruments gives inconsistent estimated coefficients if there exists white noise measurement error (Holtz-Eakin *et al.*, 1988).

External IV

This paper also uses external instruments, the head of household's income satisfaction, for the differenced lagged income and consumption. First, if external IVs are valid while internal IVs are not, the use of only external IVs presents consistent estimates. Second, the exclusion of the two-period lagged dependent variable from a set of internal IVs may cause a weak instruments problem. As only earlier lags of dependent variables are used as instruments, the relationship between an explanatory lagged dependent variable and these instruments could be relatively weaker compared to the use of the two-period lagged dependent variable as instruments. In this case, to identify measurement error avoiding a potential weak instruments problem, external IVs are additionally used for both estimates using the internal IVs that include and exclude Y_{it-2}. Finally, the external instruments are also used to test my model specification whether an ARMA(1,1) specification should be chosen instead of an AR(1) specification.

The rest of this subsection discusses the validity of the external instruments.[8] Figure 1 presents average income satisfaction by income group and by consumption group and shows that there is correlation between income satisfaction and income (or consumption). It is evident that lagged income satisfaction is strongly related to lagged income and consumption except for those who are ranked in the highest one percent of income.

However, the validity of this external instrument may be suspect especially in a level model (i.e. the model (7) or (8) , which is not first-differenced) because this variable may be correlated with time-invariant omitted variables. For instance, it can be argued that a positive mindset could be one of the omitted variables, which would be correlated with income satisfaction. First differencing, however, eliminates time-invariant unobserved heterogeneity and, thus, eliminates this problem. This study

[8]See Antman and McKenzie (2007a) for more fundamental discussion about using instrumental variables to address measurement error in estimates of earnings mobility.

also assumes a sequential exogeneity of the external instruments, which means that ε_{it}^{y} and ε_{it}^{c} are uncorrelated with the past values of the household heads' satisfaction.[9] This assumption relaxes a typical assumption that there is no time-varying unobserved heterogeneity (omitted variable).

Under the assumption, the lagged levels of this instrument must also be chosen carefully. Contrary to internal IVs, income satisfaction at any t is not correlated with the lagged random measurement error differenced term (i.e. Δv_{it-1}^{k} for k=y or c). Note that v_{it-1}^{k} is assumed to be random, and the measurement error which is likely correlated with income satisfaction is captured by mean-reverting measurement error (i.e. non-classical part of time-varying measurement error). Nevertheless, income satisfaction at t or $t-1$ may be correlated with the differenced equation error (i.e. $\Delta\varepsilon_{it}^{k}$ for k=y or c), which may represent a household income or consumption shock (Deaton, 1997). For example, one can argue that people with a negative (or positive) income shock in a particular year are more likely to report income dissatisfaction (or satisfaction). Therefore, only income satisfaction at $t-2$ and earlier are candidates for external instruments for the model which addresses measurement error. The household head's satisfaction regarding their household income is subjective, and people are likely to adapt it with the average level of income of the society; there may be no relation between income and income satisfaction over time while there is at a point in time (Easterlin, 1996). This may explain why income satisfaction in KLIPS does not vary much over time. Because of the collinearity in income satisfaction across periods, $t-4$ and earlier lags are dropped from the external instrument set.[10] However, income satisfaction at $t-3$ is still used as an instrument for a model specification test.

[9] General sequential exogeneity means that ε_{it}^{y} and ε_{it}^{c} are uncorrelated with the current or past values of the household heads' satisfaction, but this study does not assume that ε_{it}^{y} and ε_{it}^{c} are uncorrelated with the current value of the external instrument.

[10] The correlation coefficients of income satisfaction between year t and year $t-a$ are .50, .43, .37 and .35 for $a=1,2,3$ and 4 respectively.

3.3. Tests

This study examines the differences of the estimates with and without assuming random measurement error, and two subsequent tests are also conducted. For these tests, the standard error of the difference of these estimates must be calculated, as well as the difference itself. This study uses a Hausman test.[11]

The first test is to identify random measurement error. The null and alternative hypotheses are

H_0 : *There is no random measurement error in income or consumption, and*

H_1 : *There is random measurement error in the variable.*

Under the null hypothesis, two-periods lagged dependent variable is an appropriate instrument, while it is not under the alternative hypothesis. Therefore, if there is statistical difference between two estimates using different estimates based on different assumptions about the existence of random measurement error, this study rejects the assumption of no random measurement error in income (or consumption).

However, if both measurement error (v_{it}^k) and equation error (ε_{it}^k) follow a first-order moving average process in the total composite error term, the difference between the estimate of γ_y (or γ_c) using IVs that include and exclude Y_{it-2} (or C_{it-2}) cannot identify random measurement error. This is so, since $E(\ln Y_{it-2} \cdot \Delta\varepsilon_{it-1}^y) \neq 0$ and $E(\ln C_{it-2} \cdot \Delta\varepsilon_{it-1}^c) \neq 0$ as well as $E(\ln Y_{it-2} \cdot \Delta v_{it-1}^y) \neq 0$ and $E(\ln C_{it-2} \cdot \Delta v_{it-1}^c) \neq 0$. Potential biases using Y_{it-2} or C_{it-2} as instruments are thus generated not only by random measurement error but also by an MA(1) process in the equation error.

A Hausman test is, thus, now to examine AR(1) against ARMA(1,1) specifica-

[11] When the Hausman test is based on two efficient GMM estimators using two nested moment conditions, it is possible to prove that the variance of the difference of the two estimators is the difference of the variance of the estimates.

tions. The null and alternative hypotheses are

$$H_0 \quad : \quad \textit{An AR(1) specification is correct, and}$$

$$H_1 \quad : \quad \textit{An ARMA(1,1) specification is required.}$$

The set of instruments including the income satisfaction at $t-2$ and $t-3$ as well as the internal instruments $t-3$ and earlier is valid under the null hypothesis, but is invalid under the alternative hypothesis. Note that an ARMA(1,1) model includes ε^y_{it-2} (or ε^c_{it-2}) in $\Delta\varepsilon_{it-1}$, which is likely to be interpreted as income shock at $t-2$ and, thus, be correlated with income satisfaction at $t-2$. Therefore, only household income satisfaction at $t-3$ should be used as the external instrument for an ARMA(1,1) model. The estimates using two different sets of instruments are compared: one includes this external instrument at $t-2$ and the other excludes the value at $t-2$, and examine whether an AR(1) specification should be rejected in favor of an ARMA(1,1).

Furthermore, this study also examines if the measurement error, v^k_{it}, which is assumed to be random, is serially correlated given that the measurement error exists. This test can show whether the measurement error, v^k_{it}, is really random. This could be critical because the use of internal IVs, which make it possible to identify the existence of classical measurement error, could be illogical. If the measurement error is serially correlated, internal IVs are correlated with the differenced total composite error ($\lambda_k\Delta\varepsilon^k_{it} - \gamma_k\Delta v^k_{it-1} + \Delta v^k_{it}$ for k=y or c) and so violate the exogeneity to be valid instruments. For example, s-period lagged dependent variable has random measurement error at $t-s$ (i.e. v^y_{it-s} or v^c_{it-s}). If $cov(v^k_{it},\, v^k_{is}) \neq 0$ for $t \neq s$, $E((\lambda_k\Delta\varepsilon^k_{it} - \gamma_k\Delta v^k_{it-1} + \Delta v^k_{it}) \cdot v^k_{it-s}) \neq 0$ for k=y or c. However, the use of external IVs is also useful to test this using a Hausman test conditional on serially uncorrelated equation error.

The null and alternative hypotheses for a Hausman test of serial independence of measurement error, v_{it}^k, are

$$H_0 \;:\; \textit{measurement error, } v_{it}^k, \textit{ is serially uncorrelated, and}$$
$$H_1 \;:\; \textit{measurement error, } v_{it}^k, \textit{ is serially correlated.}$$

Internal IVs are valid only under the null hypothesis, but external IVs are always valid under both null and alternative hypothesis. Therefore, the comparison between two estimates used different sets of instruments indicates whether or not the measurement error, v_{it}^k, is serially correlated. One set consists of only internal IVs, and the other set consists of both internal and external IVs. In this case, two-periods lagged dependent variable is excluded from the set of internal IVs because the existence of random measurement error is assumed. If these estimates are different, the null hypothesis would be rejected.

3.4. Error Decomposition

The difference of total composite error in model (9) or (10), which is defined as $u_{it} \equiv \lambda_k \Delta\varepsilon_{it}^k - \gamma_k \Delta v_{it-1}^k + \Delta v_{it}^k$ for k=y or c, consists of equation error (i.e. ε_{it}^k for k=y or c) and random measurement error (i.e. v_{it}^k for k=y or c). The individual observation equation and random (or, now, white noise) measurement errors cannot be identified separately in this study, but the variances of the measurement error can be under some further assumptions. Specifically when a homoskedasticity assumption is added to the assumption that both errors are independent and identically distributed, the variances can be separately identified. That is,

$$\varepsilon_{it} \sim iid(0, \sigma_\epsilon^2) \;\; and \;\; v_{it} \sim iid(0, \sigma_v^2). \tag{13}$$

This assumption implies three additional moment conditions of differenced total composite error terms:

$$
\begin{aligned}
E[u_{it}u_{it}] &= 2\lambda^2\sigma_\epsilon^2 + 2(\gamma^2+\gamma+1)\sigma_v^2 &\quad (14)\\
E[u_{it}u_{it-1}] &= -\lambda^2\sigma_\epsilon^2 - (\gamma^2+2\gamma+1)\sigma_v^2 &\quad (15)\\
E[u_{it}u_{it-2}] &= \gamma\sigma_v^2. &\quad (16)
\end{aligned}
$$

Other moments conditions in addition to these three moment conditions could be found (for example, $E[u_{it}u_{it-3}]$, etc), but I only use the first three moment conditions because my data consist of relatively short time periods.

A minimum distance (MD) estimation procedure could be used to estimate $\lambda^2\sigma_\epsilon^2$ and σ_v^2 from these three overidentified equations plugging the estimated coefficient $\hat{\gamma}$ and residual $\hat{u}_{it}$.[12] However, the variances that should be non-negative can be estimated as negative by this MD estimation. This study uses a constrained MD estimator which restricts the parameter σ_ϵ^2 and σ_v^2 to be non-negative.[13] In addition, bootstrapping is used for the standard errors of the constrained MD estimates. This error decomposition serves to find how much random measurement error is spread over households and how important measurement error is.

4. Data

This study uses data from the Korean Labor and Income Panel Study (KLIPS) from 1998 to 2006. The KLIPS interviews households from Korea's seven metropolitan cities and urban areas in eight provinces. With a target of 5,000 households, 13,738 of respondents aged 15 and over were interviewed in 1998. 3,821 of the original 5,000 households

[12] The variance σ_ϵ^2 and the linear coefficient of measurement error, λ, cannot be identified seperately.
[13] The constrained MD estimation is introduced in Appendix.

were interviewed in 2006, which denotes a retention rate of about 76%, comparable to the nine waves of the Panel Study of Income Dynamics (PSID) in the U.S.[14]

Household Income Variables

I examine per capita income dynamics at the household level. I use individual labor incomes of all members in a household, and the total household income variable is generated using the labor income from individual reports, plus financial income, real estate income, and income recorded as all other source of income excluding social insurance and transfer income reported at the household level. Social insurance and transfer income are excluded, as they are basically earned through the help of others. The main object of studying income dynamics is to show the degree to which current household incomes are affected by previous household incomes in the absence of such assistance. To examine this dynamic at the per capita level, total household income is divided by household size.[15]

Substantial measurement errors in the income variable are assumed. First, a large share of my dataset may be suspected of producing recall error. There can be randomly generated recall error and also non-classical recall error related to levels of education, age and so on. In addition, self-employed individuals often generate recall error, and 37% of individuals are self-employed in KLIPS. Intentional under-reporting (or over-reporting) is also suspected. Korea has a progressive income tax system, so that households can reduce their taxable income by reporting lower amounts. Some, in the opposite case, may feel shame to report their low income status. These types of measurement error are non-classical.

KLIPS reports that around 10 percent of respondents have zero-income for each year. Though those zero-income households are different year by year, around 66 per-

[14] More detailed data description is available in Lee (2009).

[15] All income is after-tax income in units of 10,000 won ($\approx$ 9 dollars) for the year 2000.

cent of zero-income households report zero-income for two consecutive years. In addition, 85 percent of those households report positive consumption and transfer income (or social insurance), and 75 percent of heads for those households are over 60 years old. However, low-income households are important samples for this study. In other to retain these observations, per capita household income variables are altered in logarithmic forms after adding one (i.e. $ln\ (Income + 1)$).[16]

Household Consumption Variables

Like the study of income dynamics, the per capita consumption variable is constructed at the household level and is utilized as both dependent and lagged independent variables in the subsequent analysis. Though there are fewer motives to under-report in consumption, substantial recall errors are assumed because of the lack of documented records for retrospective questions. The fact that consumption is asked about aggregated groups can also lead to substantial measurement error.[17]

KLIPS reports household expenditure through two methods: through the direct reporting of total monthly household expenditure in a single question covering all consumption items and through the more common disaggregated method, which is based on details of household expenditure. However, even for the latter, KLIPS suffers from the lack of subdivision of consumption categories. Other panel surveys usually have more categories for expenditure data. Some have more than a hundred categories, but KLIPS only has 11 (for the second wave) to 20 (for the ninth wave) categories.

The survey directly asks for total household consumption for every year but ex-

[16] Using the level of income instead of taking logarithms shows a much higher standard error of the coefficients $\hat{\gamma}_y$, whereas the estimates of γ_c show robustness regardless of taking logarithms of consumption or not. The skewness of income leads to a preference in taking logarithms. Since adding one could be criticized as an arbitrary choice, additional robustness checks are conducted and shown in Lee (2009). The estimates of γ_y show robustness according to adding the numbers from 0.05 to 2.

[17] Another national representative data on consumption in Korea, the Annual Report on the Household Income and Expenditure, which resembles the US Consumer Expenditure Survey (CES), have been collected via a diary survey and with more disaggregated groups of consumption. These data certainly have less measurement error but are not longitudinal.

cludes the disaggregated details for the first and third waves. The total annual household consumption based on the direct reporting method is thus chosen for my main analysis. This is also preferred due to the lack of subdivision of consumption categories in KLIPS, which shows no difference between the aggregate and disaggregate levels of expenditure. However, a per capita household consumption variable is also constructed by aggregating subdivided consumption and its analysis is reported in the robustness section.

Unlike income variables, only two households report zero consumption. Log of per capita household consumption variables is taken without adding one; the two households who report zero consumption are excluded.

Other Control Variables

A set of household characteristics is controlled in the analyses. The set includes household size, fraction of elderly people, educational level of head of household, sex of head of household, age of the head of household and its square, a locality indicator to show whether the respondent resides in Seoul, and a non-spouse indicator to show whether the household contains a wife or husband. All control variables are treated as exogenous. The main statistics are reported in Table 1.

Income Satisfaction Variables

This paper uses the lags of household heads' measured satisfaction regarding their household income as external instruments.[18] This variable comes from the response of each household head to the question 'how much are you satisfied with your household net income' and each individual responds according to degree of satisfaction on a 1 to 5 scale, with "1" being very satisfied and "5" being very dissatisfied. Lower scores,

[18]Eating-out consumption and asset variables were also experimented as external instruments respectively for differenced income and consumption. The estimates using those are not robust, showing low F statistics in the first stage regressions.

therefore, measure higher satisfaction. In addition, unlike income or consumption variable, household income satisfaction is surveyed as current satisfaction at the time of responding to the questionnaire. Table 2 presents the number of observations in each satisfaction category. This question is asked at the individual level for each year except for the first wave. The same question is asked to each spouse of a household head, but this variable is not used as instruments because only 77 percent of those spouses responded to the question and the variation of satisfaction within a household is not large.[19]

Sample Size

The first-differenced model in most dynamics studies requires at least three years' data, whether or not measurement error is present. These three years' of panel data are enough to correct for bias due both to measurement error and unobserved heterogeneity (or omitted variables), provided the external instruments are valid. However, the Arellano-Bond or Holtz-Eakin *et al.* method, which uses internal instruments, requires at least four years' data if there is potential random measurement error.

The constructed measure of household income is available only from 1998 to 2005. The total household income variable is constructed by labor income that is asked as a current level and by other types of income that are surveyed retrospectively; labor income is available from 1998 to 2006, and the other types of income are from 1997 to 2005. Second, the total household consumption, which is directly asked, is available from 1997 to 2005 because it is surveyed retrospectively. On the other hand, each individual's household income satisfaction data is available only from 1999 to 2006. The overlapping periods for this analysis are only from 1999 to 2005. Because the Arellano and Bond method requires up to lagged $t-3$ for instruments under the assumption of

[19] The averages of responses are 3.46 and 3.41 for a head of household and his/her spouse respectively. The correlation coefficient between satisfaction of a husband and a wife is .7.

random measurement error, the income and consumption equations can only cover the years 2002, 2003, 2004 and 2005.

Although the use of two-step GMM prevents the loss of observations for the periods $t = 2002$ to 2005, the instrument Y_{it-3} or C_{it-3} plays a significant role in this study. So, sample size is restricted for the households that have the information. However, the size is further restricted because the model requires several variables including income, consumption, other explanatory variables and external instruments. A total of $11,297$ household/years are analyzed for the study of income dynamics, and a total of $11,796$ household/years are analyzed for the study of consumption dynamics.

5. Results

5.1. Main Results

The ordinary least squares (OLS) estimate without first differencing deals with neither unobserved heterogeneity nor measurement error bias, yet it is a good starting point for this study in order to provide a general idea about bias. Table 3 presents the estimates of γ_y and γ_c , which indicate the effect of past income on current income or the effect of past consumption on current consumption.[20] Using OLS without first differencing, the estimated coefficients $\hat{\gamma}_y$ and $\hat{\gamma}_c$ are .53 and .62 respectively. However, the IV method is essential if unobserved heterogeneity and measurement error in reported income or consumption are assumed. If the instrument, one-period lagged head of household's satisfaction regarding their household income, is exogenous, it would provide consistent estimates. The IV estimates in the level model for both studies of income and consumption give higher estimates $\hat{\gamma}$ (.78 and .95 respectively) than those of the OLS

[20] The main interest in this study is in γ (γ_y or γ_c) rather than other covariates. However, all other covariates have expected signs, and these are reported in Lee (2009).

results. If the potential bias is generated only by classical measurement error, then the instruments would be plausible, and the IV results would correct for the bias from classical measurement error. The rise of the estimates is consistent with this story.

However, arguably, the IV estimates in the level model are not consistent because the satisfaction variable may be correlated with both omitted variables and non-classical measurement error in the level model. One of the suspected omitted variables, for example, may be a positive or negative mindset, which is likely to be persistent over time. If people with a positive mindset are more likely to be satisfied with their income, work harder and have a greater future income, the lagged income satisfaction will be correlated with the error term. A similar argument can be applied to the study of consumption dynamics.

First differencing prevents this potential omitted variable bias provided the omitted variables are assumed to be time-invariant, but also takes away individual specific measurement error. The first-differenced models estimated using OLS give negative estimates $\hat{\gamma}$ for both studies of income and consumption, but they may be biased not only because there can still remain random measurement error but also because it includes a lagged dependent variable as an explanatory variable. To correct for such bias, two-step generalized method of moments (GMM) is used.

For the study of income dynamics, the estimate of γ_y in the first-differenced model using only external IVs, two- and three-periods lagged income satisfaction, which is consistent even without the assumption that the errors are serially uncorrelated and under random measurement error, is .55 and significant at the 1% level. With the additional use of internal IVs to these external IVs, the estimates of γ_y are .15 and .42 respectively using the internal IVs that include Y_{it-2} and exclude Y_{it-2}; both are significant at the 1% level. The estimate (.15) using internal IVs that include Y_{it-2}, however, is inconsistent if there exists measurement error. The consistent estimates that correct for potential bias from measurement error are either .55 or .42, which compared to .15 implies that random measurement error generates bias towards zero. On the other hand, these estimates (.55 or .42) are lower than the estimate (.78) of the IV without first differencing. As argued, the level IV estimates should correct for classical measurement error, but not for non-classical measurement error or omitted variables. This suggests that non-classical measurement and unobserved heterogeneity together are significant sources of upward bias.[21] Comparing the GMM to the level OLS estimates indicates that the combined effect of unobserved heterogeneity and non-classical measurement error offsets the effect of random measurement error.

The two-step GMM results for consumption dynamics show similar patterns. When only external IVs are used, the estimate of γ_c is .41. If internal IVs are additionally used, the estimated coefficients are .18 and .38 for the estimations using the internal IVs including C_{it-2} and excluding C_{it-2} respectively. Both are also significant at the 1% level. The former estimate (.18) is, however, potentially biased from classical measurement error, and it implies that again the random measurement error may generate bias towards zero. The consistent estimates (.38) which are estimated without using C_{it-2} as an IV are lower than the estimate (.95) of the IV without first differencing, suggesting again that unobserved heterogeneity and non-classical measurement error together are significant sources of upwards bias, and that the combined effect of those offsets the effect of random measurement error.[22]

The findings from the two-step GMM illustrate that random measurement error generates bias towards zero for both studies of income and consumption dynamics. However, the formal test results for the measurement error have not yet been reported.

[21] However, this study cannot statistically test this upward bias because the IVs with and without first differencing are not nested.

[22] While GMM does not use first stage regressions as does two-stage least square (2SLS), the examinations of first stage regressions are useful to check for a weak instruments problem. All corresponding first stage regressions are reported in Lee (2009).

This study uses Hausman tests to examine whether the two estimates based on different assumptions about measurement error are statistically different. Table 4 reports the results of the Hausman tests. The differences between the estimated coefficients for γ_y (or γ_c) is .28 (or .20), and the standard errors of the difference is .13 (or .07). These differences are thus statistically significant at standard levels, so that this study rejects the null hypothesis of no random measurement error for both income and consumption.

The estimates for ARMA(1,1) specifications are also presented in Table 3. The estimates $\hat{\gamma}_y$ and $\hat{\gamma}_c$ of models for ARMA(1,1) specifications are .34 and .38 respectively. These estimates are close to the estimates of models for AR(1) specifications. A formal Hausman test examines whether the estimates $\hat{\gamma}$ in an AR(1) and an ARMA(1,1) specifications are statistically different, and Table 5 reports the results. These results show that the test fails to reject the AR(1) specifications in favor of the ARMA(1,1) specifications. As mentioned, this also supports that there is random measurement error in income (or consumption), because the difference between the estimated coefficients for γ_y (or γ_c) using IVs that include and exclude Y_{it-2} (or C_{it-2}) could have been generated by an ARMA(1,1) without random measurement error.

Table 5 also reports the result of another Hausman test which tells us that the measurement error, v_{it}, in both reported income and consumption is serially uncorrelated. The use of internal IVs in addition to external IVs gives consistent estimates comparing to the estimates used only external IVs. The differences are .13 for income dynamics and .04 for consumption dynamics, and the standard errors of the differences are relatively high (.12 and .35 respectively).

5.2. Error Decomposition

This study decomposes the standard deviation of the random (more specifically, white noise) measurement error from that of the true equation error. As described in the previous section, this study uses three moment conditions of the differenced total composite error terms in a constrained minimum distance (MD) method and bootstraps to obtain the standard errors of these standard deviations.

The outcomes are presented in Table 6. However, interestingly, the standard deviation of the measurement error is much larger in the model for the income dynamics than that for consumption dynamics (.97 versus .19). This result suggests that random measurement error is more prevalent and varies more across households in the income than in the consumption data, consistent with general belief (e.g. Deaton, 1997). All these standard deviations are significant at 5% level.

This study cannot identify equation error, σ_ϵ^2 but can $(\lambda\sigma_\epsilon)^2$. However, note that previous studies (Bound and Krueger, 1991, Bound et al, 1994; Bollinger, 1998) report a point estimate for λ_y from .78 to .97 for income data.[23] Based on the studies, one might expect σ_ϵ^2 from .83 to 1.07 for income dynamics. Although studies have not yet reported the value of λ_c, if consumption data have similar patterns of income data, this suggests that the standard deviation of random measurement error is as big as the standard deviation of equation error for both income and consumption dynamics. This result supports the view that random measurement error exists in reported income and consumption and has a substantial magnitude. Interestingly, the much higher standard deviation of the equation error for the study of income compared to consumption dynamics supports the view that households smooth their consumption relative to their income in the face of shocks.

5.3. Robustness Checks

Internal IV

[23] They use the OLS method with validation data. Bound and Krueger and Bollinger use CPS-SSA matched files, and Bound et al use the PSID validation study.

In this study, the external IVs are additionally used for the purpose as mentioned, but the results for the second stage estimation using only these internal IVs are also reported in Table 7.

When only internal IVs excluding Y_{it-2} or C_{it-2} are used to investigate income or consumption dynamics, the GMM estimates of γ_y and γ_c are .34 and .38, respectively. They are significant at 1%. Hausman tests confirm the difference between the estimates with and without assuming random measurement error. The results again reject the null hypothesis of no random measurement error in income and consumption.

Durable Consumption

In the literature on consumption dynamics, durable consumption plays an important role because households are likely to postpone buying durable goods in response to an income shock (Ogaki and Reinhart, 1998, Browning and Collado, 2001). Therefore, durable expenditures are commonly excluded for the study of consumption dynamics. However, the main analysis in this study use aggregated consumption which include expenditures for durable goods. A per capita household consumption variable excluding durable expenditures can be also constructed by aggregating subdivided consumption, but this variable is only available in the second, fourth, and following waves, and hence data for the years 1997 and 1999 are lost by using this variable.

In spite of these costs, a consumption variable that excludes durable expenditures is constructed to check whether the results are sensitive to the choice of consumption variables, and the result is also reported in Table 7. The estimates are almost same as those using the direct reporting variable, and the results of Hausman test never change though it is not reported here. These robust results are not surprising because KLIPS suffers from a lack of disaggregation of consumption categories, and thus there exists little difference between the directly asked and constructed levels of expenditure.

6. Conclusion

This study emphasizes the importance of data quality, especially for reported income and consumption. This paper uses a method to examine the existence of measurement error in survey data when a direct comparison of real versus surveyed values is impossible. This method requires at least four periods of panel data. The income and consumption of lagged two periods is not a suitable internal instrument with the assumption of random measurement error for the study of income and consumption dynamics. However, the values of lagged three or more periods can be used as alternative instruments to identify the potential existence of random measurement error. This method can be useful, especially when external instruments are available to avoid a potential weak instrument problem and to check model specification.

This study uses an AR(1) model, which investigates a general dynamic relationship between current and previous per capita income (or consumption) rather the auto-correlated individual component model used in the other dynamic earnings literature. This means that a linear relationship between current and lagged income is assumed in this study, but both classical and non-classical measurement error are considered by the linear measurement error model. The econometrics allowing for both nonlinearity and non-classical measurement error would be much harder for an analysis of panel data.

Above all, my results empirically support the view that there is substantial random measurement error in reported income and consumption, which are the most popular and frequently applied variables in empirical studies. Classical measurement error is an important factor that leads to bias towards zero in the estimated coefficient of income and consumption dynamics. It is found that the combined effect of non-classical measurement error and unobserved heterogeneity leads to an upward bias in the estimates of both income and consumption dynamics, offsetting the effect of random measurement error. It is unfortunate that this analysis cannot distinguish between the effects of non-classical measurement error and unobserved heterogeneity. The inability to separately identify unobserved heterogeneity from non-classical measurement error makes it difficult to evaluate the potential effects of individual specific measurement error. One could, in an extreme case, argue that there is no effect of non-classical and that bias is completely generated by unobserved heterogeneity, or vice versa. Nonetheless, as described, there potentially exists non-classical measurement error, especially in surveyed income.

Though this study cannot also distinguish measurement error from equation error for each observation, the variance of the distribution of random measurement error is identified. This makes it possible to compare the magnitude of the variance in random measurement error for income to that for consumption. The standard deviation of random measurement error is estimated to be as large as that of the equation error for both income and consumption dynamics, suggesting that random measurement error is substantial. Interestingly, the standard deviation of the measurement error is much larger in the model for the income dynamics (.97) than that for consumption dynamics (.18), suggesting that random measurement error is more prevalent and varies more across households in income than in consumption. The standard deviation of the equation error is also potentially larger in the model for the income dynamics than that for consumption dynamics, indicating households smooth their consumption relative to their income in the face of shocks, as has been found in other studies.

These results suggest that measurement error must be considered to avoid incorrect inferences for policies regarding poverty and inequality of income. My results show that around half of current income or consumption can be accounted for by last year's income or consumption. On the other hand, the same analysis but without considering random measurement error finds that both income and consumption is much less persistent, and there seems to be more economic mobility.

References

Abowd J, Card D. 1989. On the covariances structure of earnings and hours changes. *Econometrica* 57(2): 411-45.

Ahmed N, Brzozowski M, Crossley T. 2006. Measurement errors in recall food consumption data. Institute for Fiscal Studies Working Paper W06/21.

Andrews D. 2000. Inconsistency of the Bootstrap When a Parameter Is on the Boundary of the Parameter Space. *Econometrica* 68: 399-405.

Antman F, McKenzie D. 2007a. Earnings mobility and measurement error: a pseudo-panel approach. *Economic Development and Cultural Change* 56(1):125-162.

Antman F, McKenzie D. 2007b. Poverty traps and nonlinear income dynamics with measurement error and individual heterogeneity. *Journal of Development Studies* 43(6): 1057-1083.

Arellano M, Bond S. 1991. Some tests of specification for panel data: monte carlo evidence and an application to employment equations. *Review of Economic Studies* 58(2): 277-97.

Baulch B, Hoddinott J. 2000. Economic mobility and poverty dynamics in developing countries. *Journal of Development Studies* 36(6): 1-24.

Bollinger C. 1998. Measurement error in the current population survey: a nonparametric look. *Journal of Labor Economics* 16 (3): 576-594.

Bollinger C, Chandra A. 2005. Iatrogenic specification error: a cautionary tale of cleaning data. *Journal of Labor Economics* 23(2): 235-257.

Bound J, Brown C, Mathiowetz N. 2001. Measurement error in survey data. In *Handbook of Labor Economics*, Vol. 5.

Bound, J, Krueger A. 1991. The extent of measurement error in longitudinal earnings data: do two wrongs make a right? *Journal of Labor Economics* 12: 345-68.

Browning M, Collado D. 2001. The response of expenditures to anticipated income changes: panel data estimates. *American Economic Review* 91(3): 681-92.

Browning M, Crossley T, Weber G. 2003. Asking consumption questions in general purpose surveys. *Economic Journal* 113: F540-F67.

Daniels, L. 2001. Testing alternative measures of microenterprise profits and net worth. J*ournal of International Development* 13: 599-614.

Deaton A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, MD: Johns Hopkins University Press.

Dragoset L, Fields G. 2006. U.S. Earnings mobility: comparing survey-based and administrative-based estimates. The Society for the Study of Economic Inequality Working Paper 55.

Easterlin R. 2001. Income and happiness: towards a unified theory. *The Economic Journal* 111: 465-84.

Gibson J, Glewwe P. 2005. Analysis of poverty dynamics. In *Handbook of Poverty Statistics,* Kamanou G (eds), *United Nations:* New York

Glewwe P, Hall G. 1998. Are some groups more vulnerable to macroeconomic shocks than others? Hypothesis tests based on panel from Peru. *Journal of Development Economics*, 56: 181-206.

Gottschalk P, Huynh M. 2006. Are earnings inequality and mobility overstated? The impact of non-classical measurement error." IZA Discussion Paper 2327.

Griliches Z, Hausman J. 1986. Errors in variables in panel data. *Journal of Econometrics* 31: 93-118.

Hausman J. 1978. Specification tests in econometrics. *Econometrica* 46(6): 1251-71.

Holtz-Eakin D, Newey W, Rosen H. 1988. Estimating vector autoregressions with panel data. *Econometrica* 56(6): 1371-95.

Jalan J, Ravallion M. 2002. Geographic poverty traps? A micro model of consumption growth in rural China." *Journal of Applied Economterics* 17: 329-46.

Kim B, Solon G. 2005. Implications of mean-reverting measurement error for longitudinal studies of wages and employment. *Review of Economics and Statistics* 87 (1):193-196.

Lee, N. 2009. Measurement error and its impact on estimates of income and consumption dynamics. IEPR Working Paper 08-11.

Lillard L, Willis R. 1978. Dynamic aspects of earning mobility. *Econometrica* 46(5): 985-1012.

Lokshin M, Ravallion M. 2004. Household income dynamics in two transition economies. *Studies in Nonlinear Dynamics and Econometrics* 8(3): Article 4.

MaCurdy T. 1982. The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics* 18: 83-114.

Meghir C, Pistaferri L. 2004. Income variance dynamics and heterogeneity. *Econometrica* 72(1): 1-32.

Moffit R, Gottschalk P. 1995. Trends in the covariance structure of earnings in the U.S.:1969-1987. University of Wisconsin Institute for Research on Poverty Discussion Paper 1001-93.

Morgenstern O. 1963. *On the Accuracy of Economic Observations*. Princeton, NJ: Princeton University Press.

Newhouse D. 2005. The persistence of income shocks: evidence from rural Indonesia." *Review of Development Economics* 9(3): 415-33.

Ogaki M, Reinhart C. 1998. Measuring intertemporal substitution: the role of durable goods. *Journal of Political Economy* 106(5): 1078-1098.

Pischke J. 1995. Measurement error and earnings dynamics: some estimates from the PSID validation study. *Journal of Business and Economic Statistics* 13: 305-14.

Figure 1: Average Income Satisfaction by Income Group and by Consumption Group

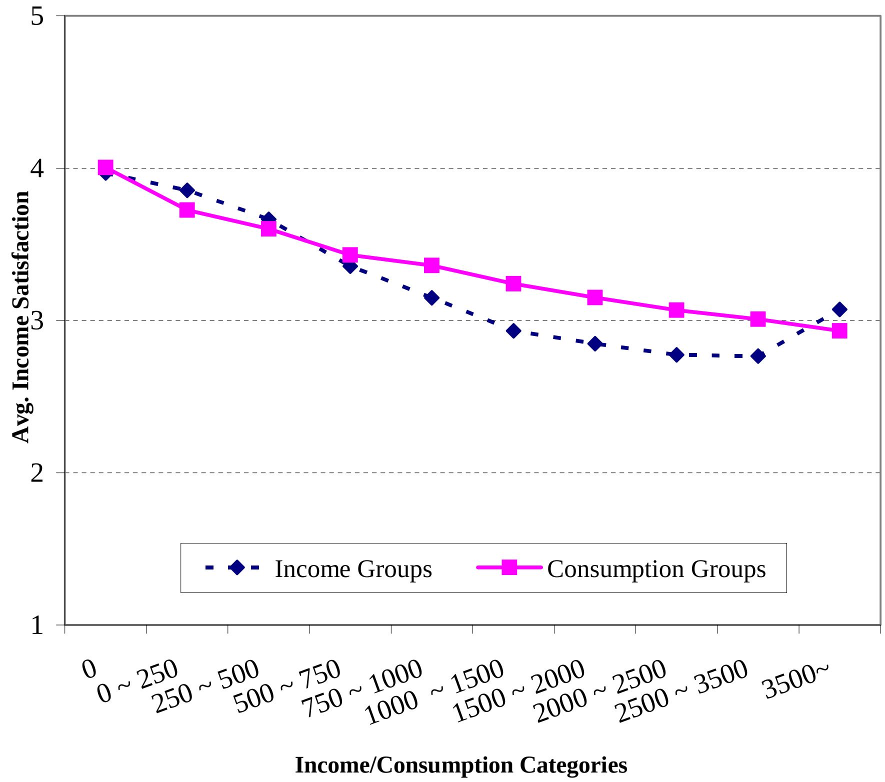

Table 1: Summary Statistics - Mean and Standard Deviation

Variable	Year 2002	2003	2004	2005	Total
For the study of income dynamics					
Log (per capita income)	5.52	5.46	5.49	5.53	5.50
	(2.24)	(2.33)	(2.34)	(2.33)	(2.31)
Household size	3.41	3.37	3.34	3.27	3.35
	(1.31)	(1.32)	(1.32)	(1.31)	(1.32)
Male aged over 65	0.06	0.06	0.07	0.08	0.07
	(0.16)	(0.17)	(0.17)	(0.18)	(0.17)
Female aged over 55	0.17	0.18	0.19	0.20	0.19
	(0.28)	(0.28)	(0.29)	(0.29)	(0.29)
Sex of head	1.16	1.16	1.17	1.17	1.17
	(0.36)	(0.37)	(0.37)	(0.38)	(0.37)
Education of head	10.18	10.20	10.23	10.27	10.22
	(4.49)	(4.47)	(4.44)	(4.44)	(4.46)
Seoul dummy	0.23	0.22	0.22	0.21	0.22
	(0.42)	(0.42)	(0.41)	(0.41)	(0.41)
Nonspouse dummy	0.21	0.22	0.22	0.24	0.22
	(0.41)	(0.42)	(0.42)	(0.43)	(0.42)
Age of head	51.41	52.82	53.73	54.56	53.14
	(12.92)	(12.84)	(12.62)	(12.53)	(12.78)
Lagged income satisfaction of head	3.52	3.42	3.44	3.46	3.46
	(0.82)	(0.78)	(0.80)	(0.80)	(0.80)
Obs #	2,803	2,756	2,871	2,849	11,279
For the study of consumption dynamics					
Log (per capita consumption)	6.06	6.12	6.13	6.17	6.12
	(0.56)	(0.56)	(0.53)	(0.54)	(0.55)
Obs #	2,930	2,902	2,974	2,990	11,796

Observations for households analyzed in this study.
Standard deviations in parentheses.

Table 2: Houshold Income Satisfaction Categories

Income	2002		2003		2004		2005		total	
Satisfaction	Obs	%	Obs	%	Obs	%	Obs	%	Obs	%
1/2	256	9.13	282	10.23	313	10.90	280	9.83	1,131	10.03
3	1143	40.78	1251	45.39	1186	41.31	1235	43.35	4,815	42.69
4	1085	38.71	1009	36.61	1144	39.85	1071	37.59	4,309	38.20
5	319	11.38	214	7.76	228	7.94	263	9.23	1,024	9.08
Total	2,803		2,756		2,871		2,849		11,279	

Lagged HH income satisfaction of head is categorized.
Observations for households analyzed in the study of income dynamics.

Table 3: Estimates of log income t-1 and log consumption t-1

Model	Estimator	Income Dynamics γ_y	Consumption Dynamics γ_c
AR(1)	*Without First Differencing:*		
	OLS	0.525	0.619
		(0.008)	(0.007)
	IVs (1)	0.782	0.954
		(0.028)	(0.027)
	After First Differencing:		
	OLS	-0.398	-0.316
		(0.014)	(0.010)
	Two Step GMM, IVs (2)	0.549	0.414
		(0.175)	(0.357)
	Two Step GMM, IVs (3)	0.145	0.172
		(0.025)	(0.018)
	Two Step GMM, IVs (4)	0.420	0.375
		(0.128)	(0.075)
ARMA(1,1)	*After First Differencing:*		
	Two Step GMM, IVs (5)	0.335	0.380
		(0.163)	(0.075)
Obs #		11,297	11,796

Robust standard errors in parentheses.
A set of household characteristics is controlled (Lee, 2009).
IVs: (1) income satisfaction of head at t-1;
(2) income satisfaction of head at t-2 and t-3;
(3) income satisfaction of head at t-2 and t-3, log income/consumption at t-2, t-3 and earlier;
(4) income satisfaction of head at t-2 and t-3, log income/consumption at t-3 and earlier;
(5) income satisfaction at t-3, and log income/consumption at t-3 and earlier.

Table 4: Hausman Tests, for time-varying measurement error

	Instruments Model A	vs	Model B	γ_A	γ_B	$\gamma_A - \gamma_B$
Income dynamics	(1)		(2)	0.420	0.145	0.275
				(0.128)	(0.025)	(0.125)
Consumption dynamics	(3)		(4)	0.375	0.172	0.204
				(0.075)	(0.018)	(0.073)

Standard errors in parentheses

IVs: (1) log income at t-3 and earlier, and income satisfaction of head at t-2 and t-3;
(2) log income at t-2, t-3 and earlier, and income satisfaction of head at t-2 and t-3;
(3) log consumption at t-3 and earlier, and income satisfaction of head at t-2 and t-3;
(4) log consumption at t-2, t-3 and earlier, and income satisfaction of head at t-2 and t-3.

Table 5: Additional Hausman Tests

	Instruments					
	Model A	vs	Model B	γ_A	γ_B	$\gamma_A - \gamma_B$
An ARMA(1,1) specification against an AR(1) specification						
Income dynamics	(2)		(1)	0.335	0.420	-0.084
				(0.163)	(0.128)	(0.100)
Consumption dynamics	(4)		(3)	0.380	0.375	0.005
				(0.075)	(0.075)	(0.008)
Serial Correlation Among Time-Varying Measurement Error						
Income dynamics	(5)		(1)	0.549	0.420	0.129
				(0.175)	(0.128)	(0.120)
Consumption dynamics	(5)		(3)	0.414	0.375	0.039
				(0.357)	(0.075)	(0.349)

Standard errors in parentheses

IVs: (1) log income at t-3 and earlier, and income satisfaction of head at t-2 and t-3;
(2) log income at t-3 and earlier, and income satisfaction of head at t-3;
(3) log consumption at t-3 and earlier, and income satisfaction of head at t-2 and t-3;
(4) log consumption at t-3 and earlier, and income satisfaction of head at t-3;
(5) income satisfaction of head at t-2 and t-3.

Table 6: Standard Deviation of Each Error Term

Type of Error	Parameter	Estimate	
		Income	Consumption
Equation Error	$\lambda\sigma_{\epsilon}$	1.066	0.233
		(.223)	(.037)
Measurement Error	σ_{v}	0.970	0.179
		(.367)	(.067)

Standard errors in parentheses

Table 7: Robustness Checks, Estimates of log income t-1 and log consumption t-1

Model	Estimator	Income Dynamics γ_y	Consumption Dynamics γ_c
Estimations using Only Internal IVs			
AR(1)	*After First Differencing:*		
	Two Step GMM,IVs (1)	0.148	0.172
		(0.025)	(0.018)
	Two Step GMM,IVs (2)	0.336	0.381
		(0.163)	(0.076)
	Obs #	11,297	11,796
Estimations for Nondurable Consumption			
AR(1)	*After First Differencing:*		
	OLS		-0.291
			(0.012)
	Two Step GMM,IVs (3)		0.512
			(0.364)
	Two Step GMM,IVs (4)		0.187
			(0.022)
	Two Step GMM,IVs (3)		0.361
			(0.076)
	Obs #		8,699

IVs: (1) log income/consumption at t-2, t-3 and earlier
(2) log income/consumption at t-3 and earlier
(3) income satisfaction of head at t-2 and t-3
(4) income satisfaction of head at t-2 and t-3, log consumption at t-2, t-3 and earlier
(5) income satisfaction of head at t-2 and t-3, log consumption at t-3 and earlier

<Not for Publication>

Appendix: Constrained Minimum Distance Estimation

This section introduces the constrained MD estimation that this study uses. To rewrite the moment conditions

$$\begin{aligned} E[u_{it}^2] &= 2\lambda^2\sigma_\epsilon^2 + 2(\gamma^2+\gamma+1)\sigma_v^2 \\ E[u_{it}u_{it-1}] &= -\lambda^2\sigma_\epsilon^2 - (\gamma^2+2\gamma+1)\sigma_v^2 \\ E[u_{it}u_{it-2}] &= \gamma\sigma_v^2 \end{aligned} \tag{17}$$

in a simple matrix form, I define

$$a = \left(E[u_{it}^2],\ E[u_{it}u_{it-1}],\ E[u_{it}u_{it-2}]\ \right)';\ b = \left(\lambda^2\sigma_\epsilon^2,\ \sigma_v^2\right)' \tag{18}$$

and

$$G(\gamma) = \begin{pmatrix} 2 & 2(\gamma^2+\gamma+1) \\ -1 & -(\gamma^2+2\gamma+1) \\ 0 & \gamma \end{pmatrix}. \tag{19}$$

Then, the moment equations can be written as

$$a = G(\gamma)\,b. \tag{20}$$

Notice that since σ_ϵ^2 and σ_v^2 are variances, the parameter b is restricted to be $b \geq 0$.

Suppose that $\hat{a}$ denotes an estimate of a. In this study, this study uses

$$\hat{a} = \left(\widehat{E[u_{it}^2]},\ E[\widehat{u_{it}u_{it-1}}],\ E[\widehat{u_{it}u_{it-2}}]\ \right)' \equiv (\hat{\omega}_0,\ \hat{\omega}_1,\ \hat{\omega}_2)', \tag{21}$$

where

$$E[\widehat{u_{it}u_{it-k}}] = \frac{1}{N(T-k)} \sum_i \sum_t \hat{u}^j_{it} \hat{u}^j_{it-k}, \ j = y, c \tag{22}$$

and

$$\hat{u}^y_{it} = \Delta \ln Y_{it} - \hat{\gamma}_y \Delta \ln Y_{it-1} - \hat{\beta}'_y \Delta X_{it} - \Delta \hat{D}_t \tag{23}$$

and

$$\hat{u}^c_{it} = \Delta \ln C_{it} - \hat{\gamma}_c \Delta \ln C_{it-1} - \hat{\beta}'_c \Delta X_{it} - \Delta \hat{D}_t. \tag{24}$$

The estimator used in this study is a constrained minimum distance (MD) estimator $\hat{b} = \left(\widehat{\lambda\sigma_\epsilon}, \hat{\sigma}_v\right)$ that solves

$$\min_{b \geq 0} (\hat{a} - \hat{G}b)' \hat{W} (\hat{a} - \hat{G}b), \tag{25}$$

where $\hat{W}$ is an estimator of the inverse of the asymptotic variance of $\hat{a}$ and $\hat{G} = G(\hat{\gamma})$.

In general, the closed form solution of the constrained optimization problem such as (25) is complicated. However, when the dimension of b is 2, the closed form solution is written in a relatively simple way; this study defines

$$\hat{V} \equiv \hat{G}' \hat{W} \hat{G} \equiv \begin{pmatrix} \hat{v}_{11} & \hat{v}_{12} \\ \hat{v}_{21} & \hat{v}_{22} \end{pmatrix} \tag{26}$$

where $\hat{W}$ is inverse of estimated covariance matrix of $\hat{a}$.

$$\begin{aligned}
\hat{W} &= \begin{pmatrix} \widehat{Var}(\hat{\omega}_0) & \widehat{Cov}(\hat{\omega}_0, \hat{\omega}_1) & \widehat{Cov}(\hat{\omega}_0, \hat{\omega}_2) \\ \widehat{Cov}(\hat{\omega}_0, \hat{\omega}_1) & \widehat{Var}(\hat{\omega}_1) & \widehat{Cov}(\hat{\omega}_1, \hat{\omega}_2) \\ \widehat{Cov}(\hat{\omega}_0, \hat{\omega}_2) & \widehat{Cov}(\hat{\omega}_1, \hat{\omega}_2) & \widehat{Var}(\hat{\omega}_2) \end{pmatrix}^{-1} \qquad (27) \\
\hat{W}_{k,l} &= E[u_{it}\widehat{u_{it-k}u_{it}}u_{it-l}] - \hat{\omega}_k\hat{\omega}_l, \\
E[u_{it}\widehat{u_{it-k}u_{it}}u_{it-l}] &= \frac{1}{N\left(T - \max\left(k, l\right)\right)} \sum_i \sum_t \hat{u}^j_{it}\hat{u}^j_{it-k}\hat{u}^j_{it}\hat{u}^j_{it-l}, \; j = y, c.
\end{aligned}$$

Let $\tilde{b}$ denote the solution of the problem (25) without the restriction of non-negative variances. That is,

$$\tilde{b} = \left(\hat{G}'\hat{W}\hat{G}\right)^{-1}\left(\hat{G}'\hat{W}\hat{a}\right). \qquad (28)$$

Then, the constrained solution of (25) solves

$$\begin{aligned}
\tilde{b}'\hat{V}\left(\tilde{b} - \hat{b}\right) &= 0 \qquad (29) \\
\hat{b} &\geq 0.
\end{aligned}$$

In this case, the constrained solution $\hat{b}$ can be interpreted as the projection of the unconstrained solution $\tilde{b}$ over the convex cone $\{b \in \mathbb{R}^2 : b \geq 0\}$, and its closed form is

$$\begin{aligned}
\hat{b} \equiv \; (\widehat{\lambda\sigma_\epsilon}, \; \hat{\sigma}_v) \; &= (\widetilde{\lambda\sigma_\epsilon}, \; \tilde{\sigma}_v) \equiv \tilde{b} & & \text{if } (\widetilde{\lambda\sigma_\epsilon})^2 \geq 0, \; \tilde{\sigma}^2_v \geq 0 \\
&= (0, \; \frac{\hat{v}_{21}}{\hat{v}_{22}}\widetilde{\lambda\sigma_\epsilon} + \tilde{\sigma}_v) & & \text{if } (\widetilde{\lambda\sigma_\epsilon})^2 < 0, \; \tilde{\sigma}^2_v \geq 0 \\
&= (\widetilde{\lambda\sigma_\epsilon} + \frac{\hat{v}_{12}}{\hat{v}_{11}}\tilde{\sigma}_v, \; 0) & & \text{if } (\widetilde{\lambda\sigma_\epsilon})^2 \geq 0, \; \tilde{\sigma}^2_v < 0 \qquad (30) \\
&= (0, \; 0) & & \text{if } (\widetilde{\lambda\sigma_\epsilon})^2 < 0, \; \tilde{\sigma}^2_v < 0.
\end{aligned}$$

For the standard error of $\hat{b} = (\widehat{\lambda\sigma_\epsilon}, \; \hat{\sigma}_v)$, bootstrapping is used.[24] For

[24] According to Andrews (2000), when the true parameter b is on the boundary of the parameter set, that is, either σ^2_ϵ or σ^2_v is zero, the conventional bootstrap estimator is inconsistent. However, the

bootstrapping, I draw non-parametric bootstrap samples 10,000 times from the time stacked household observations of $Y_i = (Y_{i1}, ..., Y_{iT})'$ and $X_i = (X_{i1}, ..., X_{iT})'$, for $i = 1, ..., N$. The parameters $\lambda^2\sigma_\epsilon^2$ and σ_v^2 are estimated using the formula of the constrained MD estimator in (30) for each boostrapped samples, say $\hat{b}^s = ((\widehat{\lambda\sigma_\epsilon})^{2,s}, \hat{\sigma}_v^{2,s})$, $s = 1, ..., 10,000$, and for the standard error of $\hat{b} = (\hat{\sigma}_\epsilon, \hat{\sigma}_v)$, the sample standard deviation of the 10,000 boostrap estimators are computed as $\frac{1}{S}\sum_{s=1}^{S}\left(\hat{b}^s - \hat{b}\right)\left(\hat{b}^s - \hat{b}\right)'$.[25] The reason not to use the standard estimator of the MD estimator's asymptotic variance is that the moment relation $G(\gamma)$ contains the unknown parameter γ, and its estimate $\hat{G} = G(\hat{\gamma})$ yields a very complicated asymptotic variance formula.

estimation results in this study convince that b is different from zero.

[25] For the study of consumption dynamics, 1,404 out of 10,000 boostrap MD estimates of $\left(\lambda^2\sigma_\epsilon^2,\ \sigma_v^2\right)$ were negative when estimation is done without the sing restriction of $\left(\lambda^2\sigma_\epsilon^2,\ \sigma_v^2\right) \geq 0$. In the case of income, there were 1,744 negative unconstrained bootstrap MD estimates out of 10,000 bootstrap repetitions.