THE PHILIPPINE STATISTICIAN

An Official Publication of the
Philippine Statistical Association Inc.

November 2008, Volume 57

The Philippine Statistician is the official scientific journal of the Philippine Statistical Association, Inc.

The Editorial Board acknowledges the invaluable contribution of the following reviewers:

Jose Ramon Albert, Ph. D.	Philippine Institute for Development Studies
Mercedes B. Concepcion, Ph. D.	National Academy of Sciences and Technology
Charu Garg, Ph. D.	The World Bank
Guntur Sugiyarto, Ph. D.	Asian Development Bank
Ajay Tandon, Ph. D.	The World Bank

THE PHILIPPINE STATISTICIAN

November 2008, Volume 57

An official publication of the Philippine Statistical Association Inc. 2008

Contents

The Philippine Statistician
Vol. 57, 2008 © Philippine Statistical Association, Inc.
ISSN 2094-0343

Identifying the Determinants of Overseas Filipinos' Remittances: Which Exchange Rate Measure is Most Relevant?

Francisco G. Dakila, Jr. and Racquel A. Claveria[1]

Abstract

The study considers which exchange rate measure is the most relevant determinant of overseas Filipinos' (OF) remittances. The study employs a microeconomic framework (involving a utility maximizing household) to help form the basis for selection of variables to be included in the estimations. Different versions of the vector autoregressions (VARs), which vary by the particular exchange rate measure utilized, were estimated to quantify the impacts of the major influences on OF remittances. An OF deployment-based effective exchange rate index was constructed in the study which proved to be a significant predictor of the movement of OF remittances. Moreover, an incidental but significant finding of the study is the procyclical nature of OF remittances.

Keywords: Cyclical behavior, index numbers, exchange rate, vector autoregression, Granger causality test.

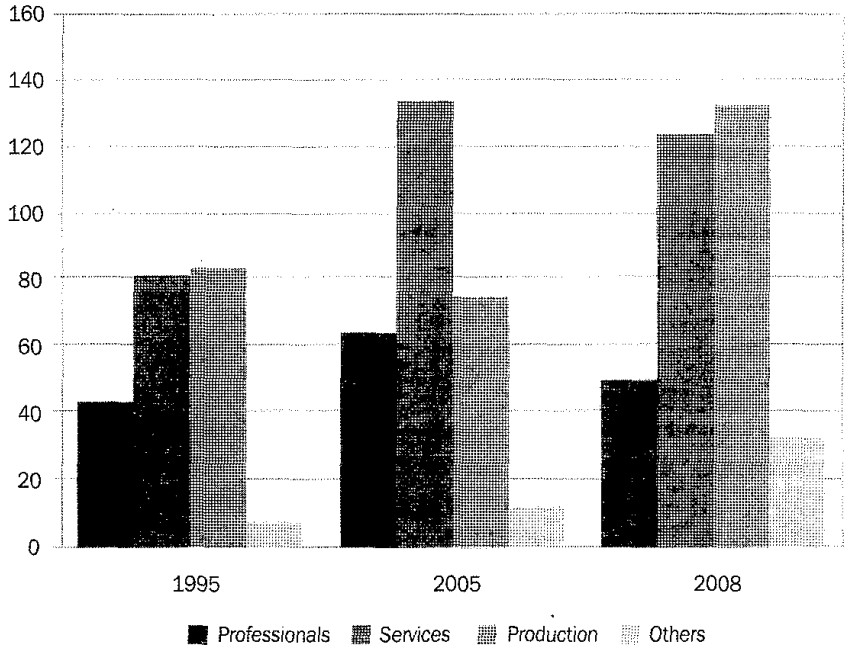
1. Introduction

Remittance inflows have increased substantially as the stock of overseas Filipinos (OFs) has grown and shifted towards more skilled jobs (Figure 1).[2] Aside from exports of goods and services, remittances have become the largest foreign exchange source for the Philippines. Remittances have been a relatively stable source of foreign exchange compared to foreign direct investment (FDI) and other private capital inflows (Figure 2). Geographic diversification of OFs may also have contributed to this stability. Globally, the Philippines' rank in the roster of top remittance-recipient countries climbed to the fourth place in 2006 from fifth in 2005 (Figure 3).

1 The authors are, respectively, Director and Bank Officer V, of the Center for Monetary and Financial Policy, Bangko Sentral ng Pilipinas (BSP), with e-mail addresses fdakila@bsp.gov.ph and rclaveria@bsp.gov.ph. The views expressed are those of the authors and do not necessarily reflect the official stance of the BSP.

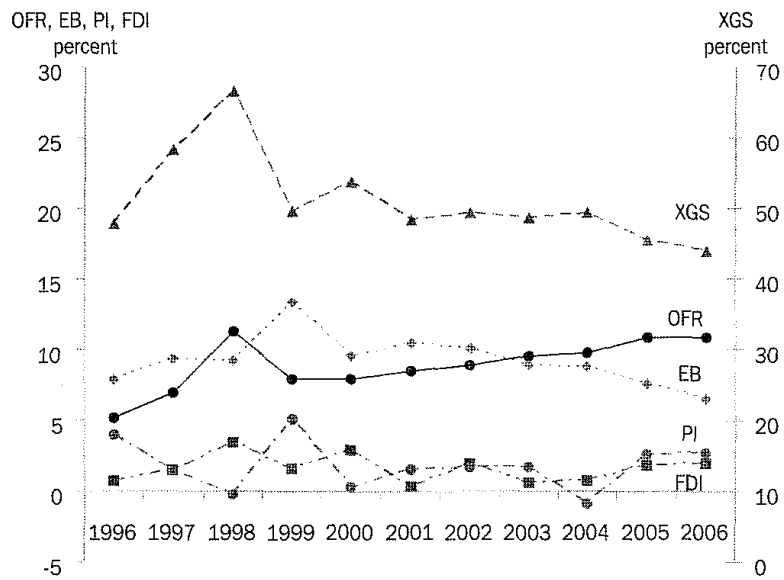
2 The Commission on Filipinos Overseas (CFO) classifies overseas Filipinos (OFs) into three categories, namely: permanent immigrants or legal permanent residents abroad whose stay do not depend on work contracts; temporary or persons whose stay overseas is related to job contracts; and irregular or those without valid residence or work permits, or those who are overstaying in a foreign country.

Figure 1 **OFW Deployment by Skill Category (in thousands)**



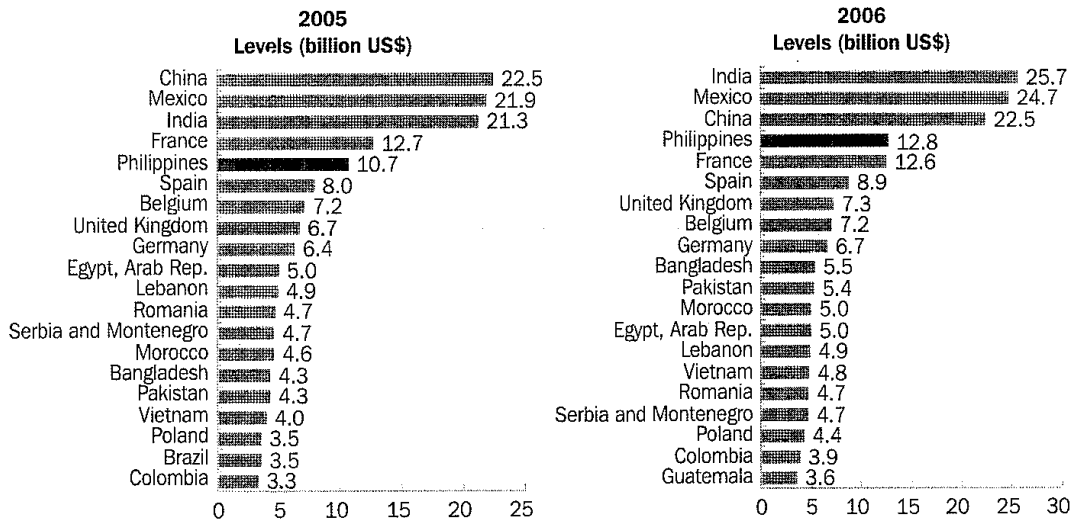
Source: BSP Department of Economic Statistics.

Figure 2 **Sources of Selected Foreign Exchange Inflows[1/] (as percentage of GDP)**



[1/]XGS – Exports of Goods and Services OFR – OF Remittances
EB – External Borrowing PI - Portfolio Investments FDI – Foreign Direct Investments
Source: BSP Department of Economic Statistics, authors' calculations.

Figure 3 **Top 10 Remittance-Remittance Recipient Countries (in billion US$)**



Sources: World Bank Global Economic Prospects 2006: Economic Implications of Remittances and Migration Migration and Remittances Factbook <http://www.worldbank.org/prospects/migrationandremittances>.

Given OF remittances' increasing trend and growing importance to the Philippine economy, their determinants, apart from their macroeconomic implications, merit greater attention. However, it would be worthwhile to distinguish first the determinants of remittances on the theoretical level (i.e., the motivations of migrant workers to remit) and on the empirical level (i.e., the measurable microeconomic and macroeconomic factors affecting the OFs' decision to remit).

On the theoretical level, altruism and self-interest constitute the two opposite ends in the spectrum of motives of migrant workers to remit income earned overseas.[3] In general, altruistic motives prevail when a decline in the recipients' income leads to a rise in remittances while self-interest or profit-driven motives dominate when an increase in the recipients' income induces an increase in remittances. Recent literature on remittances has paid greater attention to the investment and portfolio allocation considerations of migrants to remit. Under this approach, migrants decide how large a share of their incomes should be remitted to their home country to be invested there. As a consequence, remittance flows tend to respond positively to improvements in home country conditions when remittances are channeled to investments in real estate, business and capital goods.

3 In between the two extreme motives are what Lucas and Stark (1985) termed as "tempered altruism" and "enlightened self-interest".

On the empirical level, the literature on remittances tends to focus broadly on three sets of macroeconomic determinants. These are: (1) those associated with the recipient country (e.g., output, stock and skill level of migrants, unemployment and interest rates, measure of income inequality and the size of the informal economy in the home country); (2) those related to the host country (e.g., output, inflation and interest rates); and (3) those pertaining to the relationship between the host and recipient country, along with factors affecting both countries simultaneously (e.g., bilateral exchange rate, inflation differential, bilateral export flows, indicator for the development of the financial nexus between the host and home countries).

In this paper, we take a closer look at the determinants of OF remittances on the empirical level to gain understanding of what motivates OFs' decision to remit at the theoretical level. We focus on the third set of empirical determinants of remittances, particularly the exchange rate. A novel contribution of this paper is the construction of an effective exchange rate index based on the deployment of OFs. Such exchange rate index proves to be a significant predictor of the movement of OF remittances and more importantly, shows that OFs are induced by profit-driven (particularly investment) rather than altruistic motives. These results square well with the incidental finding of this paper that remittances are procyclical, disputing previous claims that OF remittances are compensatory in nature. Hence, this paper in effect contributes to two strands in the literature, i.e. the impact of the exchange rate movements on remittances and the procyclicality of remittances.

The structure of the paper is as follows: Section 2 reviews the literature on the impact of exchange rate changes on remittances and the relationship between growth and remittances, focusing on the Philippine case. Section 3 develops a theoretical framework that underlies the relationship between the exchange rate and remittances. Section 4 shows the estimation methodology employed in the study with a special focus on the various exchange rate measures compiled by the Bangko Sentral ng Pilipinas (BSP). Section 5 presents the estimation results. Section 6 introduces an exchange rate index based on OF deployment and shows that such index is a significant predictor of remittances. Section 7 concludes and offers some policy prescriptions.

2. Review of Literature

On the whole, a comparatively small proportion of research has incorporated formal econometric estimates of the relationship between the exchange rate and remittances, despite the exchange rate being probably the most important link between the overseas worker and the recipient families. With regard to

the Philippines, only a couple of studies (Yang 2006, Leuth and Ruiz-Arranz 2006) included the impact of exchange rate changes on OF remittances in their empirical investigation. Yang (2006) found that OF remittances tend to increase when the peso depreciates relative to currencies of OFs' destination countries. Meanwhile, the panel estimation of Leuth and Ruiz-Arranz (2006) for 11 developing countries including the Philippines, showed that depreciation of the home country's currency reduces remittances as less amount of the host country's currency could buy the same basket of goods as before the depreciation. For other studies that covered several Asian, Caribbean and Latin American countries, the results are likewise varied. Some found that a real depreciation leads to an increase in remittances (Faini 1994, Loser et al. 2006[4]) while others showed that a real appreciation prompts a rise in remittance flows (e.g., Leuth and Ruiz-Arranz 2006). Meanwhile, several researches yielded insignificant impact of exchange rate movements on remittances (Higgins 2004, Faini 2006, Vargas-Silva 2007).

On the other hand, empirical studies on the procyclicality of remittances that included the Philippines in the estimations are likewise sparse. Noteworthy is the study of Chami, et al. (2003) that employed a sample of 113 countries (including the Philippines) for the years 1970-1998 in their panel estimation. Chami, et al. found that remittances are countercyclical (i.e., tend to rise during economic downturns) and compensatory (i.e., tend to make up for bad economic outcomes) and that motives for remitting are predominantly altruistic. In order to validate the main findings of the Chami, et al. (2003) study and gauge their applicability to the Philippine situation, the BSP (2003) reestimated the main equations of the study using Philippine data. The results of the BSP's replication of Chami, et al. (2003) indicated that while a negative relationship between remittances and growth appears in OLS estimates for the Philippines, this relationship vanishes when the appropriate correction is made for serial correlation. The divergent results obtained by Chami, et al. (2003) and BSP (2003) may indicate a need to exercise greater care in applying the results from a panel estimation of different countries to all the countries included in the panel, because of possible heterogeneity in the characteristics of the sample included in the study. Burgess and Haksar (2005) have also reexamined the link between remittances and growth, using Philippine data specifically. As with the BSP study, they did not find empirical support for the hypothesis that remittance flows exert a short-term stabilizing effect on consumption. Moreover, the authors noted that measurement issues, as well as endogeneity of regressors and the resulting problem of finding adequate instruments, can complicate the estimation of the remittances-growth relationship using macroeconomic data. In addition, Tuaño-Amador, et al.

4 For a subsample of countries covered by the study that included Columbia and El Salvador.

(2007) analyzed the cyclical components of GDP and OF remittances and found that OF remittances tend to be procyclical. Interestingly, the finding of Tuaño-Amador, et al. (2007) on the procyclicality of OF remittances using pairwise correlation techniques is supported by the validation exercise that they also performed on the study of Chami, et al. (2003) These results are, therefore, in line with the previous caution on generalizing the results obtained from a panel of countries.

3. The link between the exchange rate and of remittances: theoretical framework

We consider a representative household that maximizes the present value of its expected utility stream, defined as

$$U = E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{1}$$

where $0 < \beta < 1$ is a subjective rate of discount, c_t represents total household consumption at time period t, and u is the instantaneous utility function, assumed to be a strictly increasing and nonnegative function of c_t. The household has a fixed labor endowment L at each point in time, and derives income from supplying some amount

$$l_t \leq L \tag{2}$$

of such endowment to the labor market.

Given a wage rate w_t, the household's income level y_t is

$$y_t = w_t l_t \tag{3}$$

and at any moment of time, the household's asset accumulation is described by the relationship

$$A_{t+1} = (1 + r_t)(A_t - c_t) + y_{t+1} \tag{4}$$

where r_t is some measure of the relevant interest rate that represents the returns to the household from allocating some of its resources to asset accumulation. Combining this with the previous equation yields the following sequence of budget constraints for the household:

$$c_t + \frac{A_{t+1}}{1 + r_t} = A_t + \frac{w_{t+1} l_{t+1}}{1 + r_t} \tag{5}$$

We extend this basic framework to the decision processes of OF households by observing the following essential difference: that an OF household supplies labor to a factor market in which the remuneration is denominated in a foreign currency. Suppose that there are two labor markets open to the household–a domestic market and a foreign market, to which the household supplies amounts l_t and l_t^*, respectively. The labor supply constraint (2) should now be modified to

$$l_t + l_t^* \leq L \tag{6}$$

Given this view, we now think of some measure of the foreign wage rate, w_t^*, and modify (3) to

$$y_t = w_t l_t + e_t w_t^* l_t^* \tag{7}$$

To sum up, the OF household's utility maximization problem can now be described as

Maximize

$$U = E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{8}$$

subject to

$$c_t + \frac{A_{t+1}}{1+r_t} = A_t + \frac{w_{t+1} l_{t+1} + e_{t+1} w_{t+1}^* l_{t+1}^*}{1+r_t}$$

$l_t + l_t^* \leq L$, and

$c_t \geq 0$

This yields a sequence of vectors (c_t, l_t, l_t^*), $t = 1, 2, \ldots, n$, that are functions of current and expected values of the domestic and foreign wage rates, the exchange rate and rates of interest $(w_t, w_{t+1}, \ldots, w_t^*, w_{t+1}^*, \ldots, e_t, e_{t+1}, \ldots, r_t, r_{t+1}, \ldots)$. Thus, at this point, we note that the forward-looking analysis injects several sources of uncertainty into the framework. In particular, these pertain to the wage rate, the interest rate and the exchange rate. The wage rate is typically subject to longer-term contractual arrangements, and is relatively more predictable. On the other hand, while the implications of variability of the interest rate on household asset accumulation are an important issue, especially in terms of the impact on the monetary transmission mechanism, the issue is not unique to OF households. In this paper, therefore, we focus on the link between the exchange rate and remittances. Note that the specification of the household's budget constraint assumes that foreign wage earnings enter into the asset accumulation

equation using the exchange rate at the time that the foreign compensation was earned. With the variety of exchange rate measures in the market, exchange rate considerations influence the migrant's decision to remit. The majority of previous studies that analyzed the impact of exchange rate changes on remittances (as discussed in Section 2) employed real bilateral exchange rates as the relevant exchange rate measure. However, in view of the foregoing discussion, using such measure of exchange rate could lead to measurement error and biased estimates. Hence, it would be worthwhile to consider an exchange rate measure that compares the migrant's relative earnings opportunities in the home vis-à-vis the destination countries.

As shown in Section 2, the various pieces of empirical evidence on the relationship between the exchange rate and remittances lend support to the notion that the impact of exchange rate changes on remittances is a priori ambiguous. Theoretical discussions in the literature generally show that a real depreciation in the home country could either lead to a rise or fall in remittances sent by overseas workers depending on their motives to remit and on their labor supply behavior. A depreciation of the currency in the home country constitutes a positive income shock to the recipients of remittances since each unit of foreign currency would be converted to more units of local currency once remitted. If the migrant worker is driven by altruistic motives such that the purpose of the transfer is to meet the required expenses of the household, then the migrant will reduce the amount of foreign currency that he/she is sending back home. On the other hand, if the overseas worker is motivated by profits or the desire to finance an investment in the home country, then the migrant could take advantage of the depreciation by sending more remittances to the home country.

However, when the labor supply behavior of migrants becomes intertwined with the motives to remit, the impact of exchange rate changes on remittances is not as clear-cut. The labor supply behavior of migrant workers likewise reacts to exchange rate movements via the impact of the latter on the income level of the migrant when expressed in home currency. The depreciation of the home currency increases the migrant's wage value when denominated in the home currency. With higher wages, the migrant could either increase his/her labor supply as leisure becomes relatively expensive (substitution effect) and hence remit more. Alternatively, the higher value of the migrant's remuneration brought forth by the depreciation of the home currency could induce him/her to decrease labor supply (i.e., prefer leisure from work) once the recipients' estimated need of the migrants' wage income is satisfied (income effect) and therefore remit less. Hence, remittances could either rise or fall depending on whether the substitution or income effect is dominant. The dominance of the income effect (i.e., the desire to maintain remittances' purchasing power) is more easily reconcilable with altruistic behavior than with a quest for profits. Meanwhile, the dominance of the substitution effect

(i.e., the desire to take advantage of the depreciation by remitting more) is deemed more consistent with profit-driven (most plausibly investment) motives of OFs.

4. Data and estimation methodology

Our framework forms the basis for our selection of variables to be included in our equation system. In order to quantify the impacts of the major influences on OF remittances, we utilized vector autoregressions (VARs) estimated using quarterly data of the following variables: real 91-day T-bill rate, exchange rate measure, gross domestic product (GDP) per capita, OF deployment per capita, and OF remittances per capita, for the sample period 1980-2005. In the absence of appropriate wage rate data for OF households in the Philippines, we used GDP as a proxy for the wage rate as an indicator of earnings opportunities. In this respect, it can be noted that variation in the level of GDP can be expected to induce substitution and income effects on the level of remittances, in the same manner as variation in the wage rate. While it would be an interesting exercise to compile an index of earnings opportunities for countries in which OFs are deployed, including this in the VAR would have implications on the degrees of freedom, given the available data sample.

Meanwhile, OF remittances data compiled by the BSP were utilized in this study. The BSP applies a raising factor (derived from the Survey of Filipinos, a rider to the Labor Force Survey of the National Statistical Office) to cash remittances coursed through the banking system to obtain global cash remittances (i.e., flows coursed through formal and informal channels). On the other hand, the data series for deployment of OFs based on countries of destination are sourced from the Commission on Filipino Overseas (CFO) and Department of Foreign Affairs (DFA) and are available from 1996 onwards.

Several different versions of the VARs were estimated, which vary by the particular exchange rate measure utilized. As a first step, we specified the VAR using the exchange rate measures and indices regularly compiled by the BSP, namely, the peso-dollar exchange rate (nominal and real bilateral) and six trade-weighted exchange rate indices computed by the BSP (three nominal effective exchange rate (NEER) measures, viz., NEER broad, NEER narrow, and NEER major, corresponding to different currency baskets, each of which is then adjusted for inflation differentials to come up with three corresponding real effective exchange rate (REER) measures, viz., REER broad, REER narrow, and REER major). Box 1 describes the effective exchange rate indices compiled by the BSP in greater detail.

Box 1 **Effective Exchange Rate Indices of the BSP**

The Nominal Effective Exchange Rate (NEER) indices for the peso are the weighted average exchange rate of the peso vis-à-vis a suitably chosen basket of foreign currencies, unadjusted for the effects of inflation. This is computed as the summation of the percentage changes in the peso cross rates with respect to each currency in the basket multiplied by the corresponding country weight. The individual country weights, in turn, are calculated from their total trade shares, i.e., exports plus imports.

The Real Effective Exchange Rate (REER) indices for the peso are the corresponding NEER indices of the peso adjusted for inflation rate differentials with the countries whose currencies comprise the NEER basket. This is simply the peso's NEER index multiplied by the ratio of the domestic price index to the weighted price index of the countries whose currencies comprise the NEER basket. The base year used for the consumer price indices of the Philippines and countries included in currency baskets is 1980.

The NEER and REER indices are computed for three groups of countries as summarized in the following table:

Box Table 1 **Composition of Currency Baskets**

Country	Narrow	Broad	Major
United States			✓
Japan			✓
European Monetary Union (EMU)			✓
United Kingdom			✓
Singapore		✓	
South Korea		✓	
Taiwan		✓	
Malaysia	✓	✓	
Thailand	✓	✓	
Indonesia	✓	✓	

5. Estimation Results

As a guide to the appropriate ordering of variables in the VAR, pair wise Granger causality tests were performed. The results (Table 1) indicate that the chain of (Granger) causation runs from remittances to the exchange rate, and then on to the real interest rate, and finally, to GDP. There is some bi-directional causality between remittances and OF deployment. The VAR ordering broadly follows these results, but we have placed deployment as the last variable in the VAR, since we expect the decision to work abroad as responding largely to longer-term considerations.

Our initial assessment of the relevance of the various exchange rate measures in explaining the magnitude of remittances considers the response of the latter to a one-standard deviation shock in the former, as indicated in the impulse response function (IRF). The nominal and real peso-dollar rate, plus six other real and nominal trade-weighted exchange rate indices were tested, namely, the real effective and nominal effective exchange rates for the broad, narrow and major-trading partners basket of currencies. The

Table 1. Pairwise Granger Causality Tests

Sample: 1980Q1 200504	Lags: 4		
Null Hypothesis:	**Obs**	**F-Statistic**	**Probability**
LOG(EXCHANGE RATE) does not Granger Cause LOG(OF REMITTANCES)	64	0.28	0.89
LOG(OF REMITTANCES) does not Granger Cause LOG(EXCHANGE RATE)		2.87	0.03
REAL 91-DAY T-BILL RATE does not Granger Cause LOG(OF REMITTANCES)	64	1.27	0.29
LOG(OF REMITTANCES) does not Granger Cause REAL 91-DAY T-BILL		2.17	0.08
LOG(GDP) does not Granger Cause LOG(OF REMITTANCES)	64	0.52	0.72
LOG(OF REMITTANCES) does not Granger Cause LOG(GDP)		0.15	0.96
LOG(OF DEPLOYMENT) does not Granger Cause LOG(OF REMITTANCES)	60	3.08	0.02
LOG(OF REMITTANCES) does not Granger Cause LOG(OF DEPLOYMENT)		2.51	0.05
REAL 91-DAY T-BILL RATE does not Granger Cause LOG(EXCHANGE RATE)	100	0.13	0.97
LOG(EXCHANGE RATE) does not Granger Cause REAL 91-DAY T-BILL RATE		4.47	0.00
LOG(GDP) does not Granger Cause LOG(EXCHANGE RATE)	96	2.17	0.08
LOG(EXCHANGE RATE) does not Granger Cause LOG(GDP)		15.31	0.00
LOG(OF DEPLOYMENT) does not Granger Cause LOG(EXCHANGE RATE)	96	1.08	0.37
LOG(EXCHANGE RATE) does not Granger Cause LOG(OF DEPLOYMENT)		0.83	0.51
LOG(GDP) does not Granger Cause REAL 91-DAY T-BILL RATE	96	0.63	0.65
REAL 91-DAY T-BILL RATE does not Granger Cause LOG(GDP)		12.65	0.00
LOG(OF DEPLOYMENT) does not Granger Cause REAL 91-DAY T-BILL RATE	96	0.41	0.80
REAL 91-DAY T-BILL RATE does not Granger Cause LOG(OF DEPLOYMENT)		2.48	0.05
LOG(OF DEPLOYMENT) does not Granger Cause LOG(GDP)	92	0.97	0.43
LOG(GDP) does not Granger Cause LOG(OF DEPLOYMENT)		1.84	0.13

Note: Granger causality tests employ the Granger approach to the question of whether X causes Y by determining how much of the current Y can be explained by past values of Y and then seeing whether adding lagged values of X can improve the explanation of Y. Y is said to be Granger-caused by X if X helps in the prediction of Y, or equivalently if the coefficients on the lagged X's are statistically significant.

Source: Authors' calculations based on data from the BSP Department of Economic Statistics.

Table 2. Specification of Estimated Vector Autoregressions (VARs)

Variable	VAR1	VAR 2	VAR 3	VAR 4	VAR 5	VAR 6	VAR 7	VAR 8
Log(OF Remittances)	✓	✓	✓	✓	✓	✓	✓	✓
Exchange Rate (P/$)	✓							
Real Exchange Rate ($/P)		✓						
NEER Broad			✓					
REER Broad				✓				
NEER Narrow					✓			
REER Narrow						✓		
NEER Major							✓	
REER Major								✓
Real 91-day T-bill Rate	✓	✓	✓	✓	✓	✓	✓	✓
Log(GDP)	✓	✓	✓	✓	✓	✓	✓	✓
Log(OF Deployment)	✓	✓	✓	✓	✓	✓	✓	✓

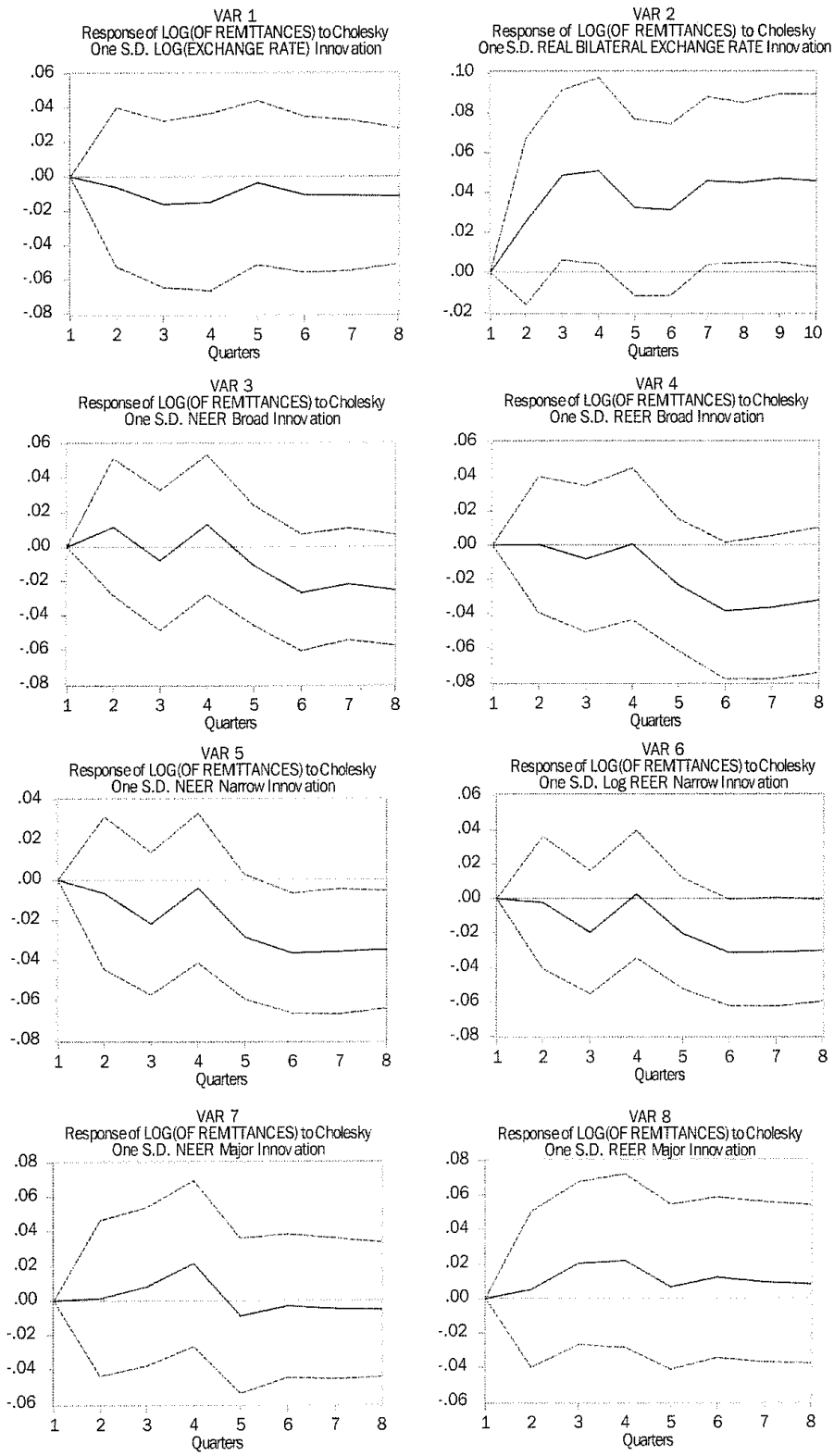
Source: Authors' calculation based on data from the BSP Department of Economic Statistics.

composition of the various estimated VARs are summarized in Table 2. The impulse responses corresponding to the different VARs are presented in Figure 4, and can be summarized as follows:

- **Response to nominal peso-dollar rate.** As discussed in Section II, a depreciation of the currency implies greater purchasing power for each dollar of remittance, and should therefore encourage remittances. This constitutes a substitution effect, and implies positive values for the IRF of remittances to both the nominal and real peso-dollar rate. On the other hand, it may happen that the OF families' expenditure baskets include some fixed, recurrent items that are denominated in pesos, including education and rentals, for example. In this case, a peso depreciation tends to reduce remittances, so that, as with most other price changes, the direction of this income effect is opposite to that of the substitution effect. In most cases, the substitution effect can be expected to dominate the income effect. However, the estimated responses to the nominal peso-dollar rate from the IRF are statistically insignificant and generally of the opposite direction from the expected response. It can be noted that the insignificant result is in accordance with the Granger causality tests presented in Table 1.

- **Response to real peso-dollar rate**. It is possible that the insignificant result from the preceding estimates is due to the failure of the NER to account for differences in the purchasing power of the peso vis-à-vis foreign currencies. As a first step, we examined the IRF from a VAR incorporating a real bilateral exchange rate measure (US$/₱), which was defined such that an increase in the measure represents an appreciation of the peso against the US dollar, in order to align this measure with the other exchange rate indices computed by the BSP. In contrast to the results for the NEER, the IRF for the real bilateral rate shows a statistically significant response of remittances. However, contrary to expectations (i.e., assuming dominance of the substitution over the income effect), an appreciation of the peso induces an increase in remittance flows, which is quite sustained over the ten-quarter simulation period. The results indicate that the motive to finance expenditures that are fixed in peso terms can be dominant among OF families. Moreover, the results also run counter to the widely-held belief that remittances are countercyclical. In fact, if weakening of the peso is associated with downtrends in the economy, and vice-versa, then the results can indicate a potential for remittance flows to actually

Figure 4 **Impulse Responses of OF Remittances from VARs 1-8**



Note: In a VAR system, a shock to the i-th variable not only directly affects the i-th variable but is also transmitted to all of the other endogenous variables through the dynamic (lag) structure of the VAR. Impulse response functions trace the effect of a one-time shock to one of the innovations on current and future values of the endogenous variables. The transformation of impulses in the impulse response functions above employs the inverse of the Cholesky factor of the residual covariance matrix to orthogonalize the impulses. This imposes an ordering of the variables in the VAR and attributes all of the effect of any common component to the variable that comes first in the VAR system.

Source: Authors' calculations based on data from the BSP Department of Economic Statistics.

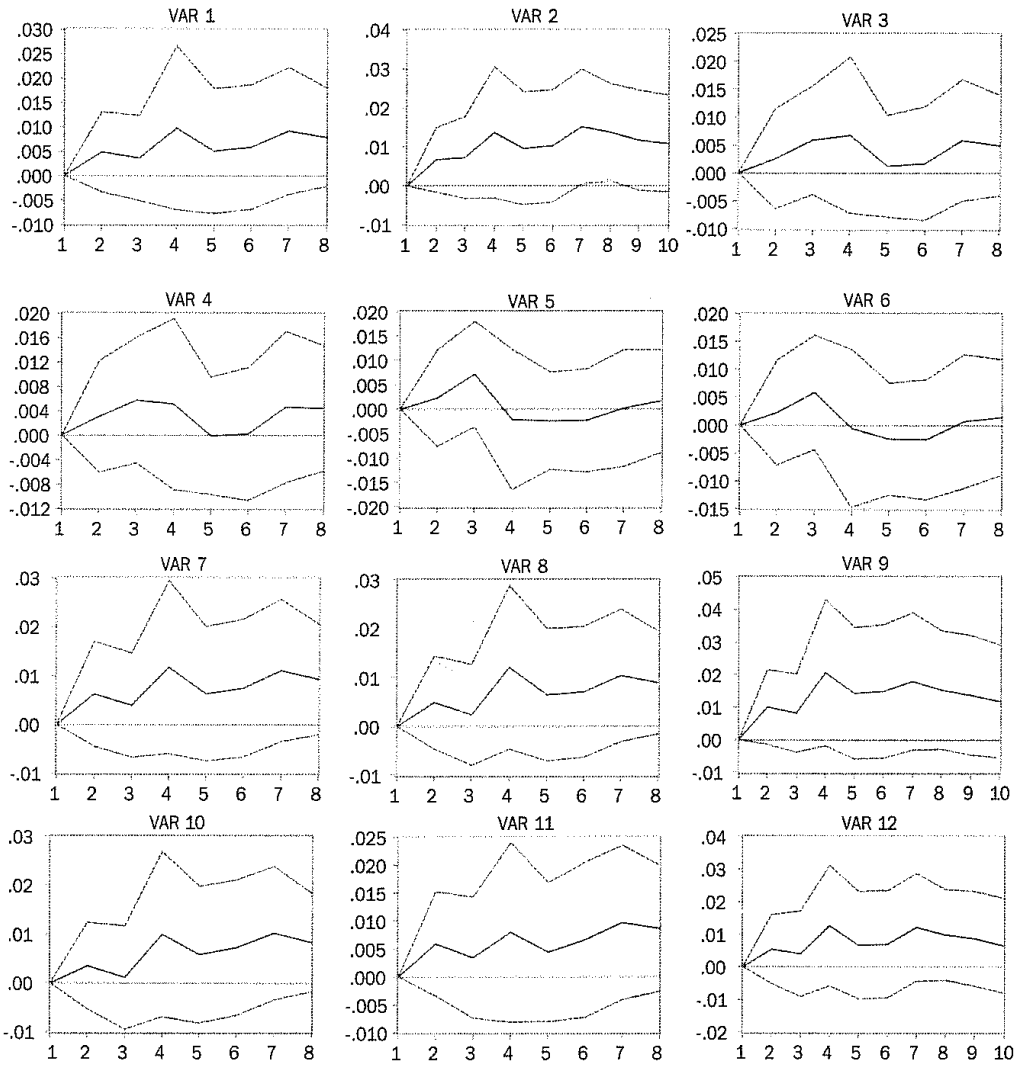
exacerbate the cyclicality of economic activity. We also note that, for both the nominal and real peso-dollar rate, the results support dominance of the income effect over the substitution effect.

- **Response to various nominal and real exchange rate measures.** All the nominal (NEER) and real (REER) effective exchange rate indices are defined so that an increase in the index represents an appreciation of the peso against the basket of currencies comprising the index. Broadly significant responses are obtained for the real and nominal indices for the "Broad" and "Narrow" groups of countries. This contrasts to the insignificant results obtained for the "Major" country grouping, which consists of some developed country trading partners of the Philippines. Moreover, in contrast to the results for the bilateral exchange rates, the impulse responses for the foregoing indices are in the "expected" direction (i.e., with the substitution effect dominating the income effect). This points out further the need to explain the variance of the results for the bilateral exchange rates from the "expected" response. In brief, exchange rate indices that measure the "expensiveness" of domestic products vis-à-vis a basket of products of competitor countries appear to elicit the expected substitution effect, in contrast to movements in the peso-dollar rate. We return to this issue in Section 5.

Are Remittances Pro- Or Anti-cyclical?

Although the matter of the pro- or countercyclicality of OF remittances is not the main topic of this paper, it can reinforce our previous finding that motives of OFs are predominantly profit-driven or investment-related as seen from the positive impact on remittances of a peso depreciation vis-à-vis currencies of top OF host countries. As we have indicated in Section II, profit-driven motives are apparent when an increase in the recipients' income encourages OFs to remit more. Figure 5 collates the IRFs from the previously estimated VARs, pertaining to the response of OF remittances to a shock in GDP. Although most of the results are statistically insignificant, the general direction of the responses is consistent across the VAR specifications: a positive GDP shock induces an increase in remittances. Statistically significant results are obtained when the real bilateral (₱/$) exchange rate measure is included in the VAR. These results indicate that OF remittances tend to be procyclical and that motives of OFs to remit are likely to be profit-driven rather than altruistic.

Figure 5 **Impulse Responses of OF Remittances to GDP from VARs 1-12 (Response of LOG(OF REMITTANCES) to Cholesky One S.D. LOG (GDP) Innovation)**



Source: Authors' calculations based on data from the BSP Department of Economic Statistics.

6. An Exchange Rate Measure Based On Of Deployment

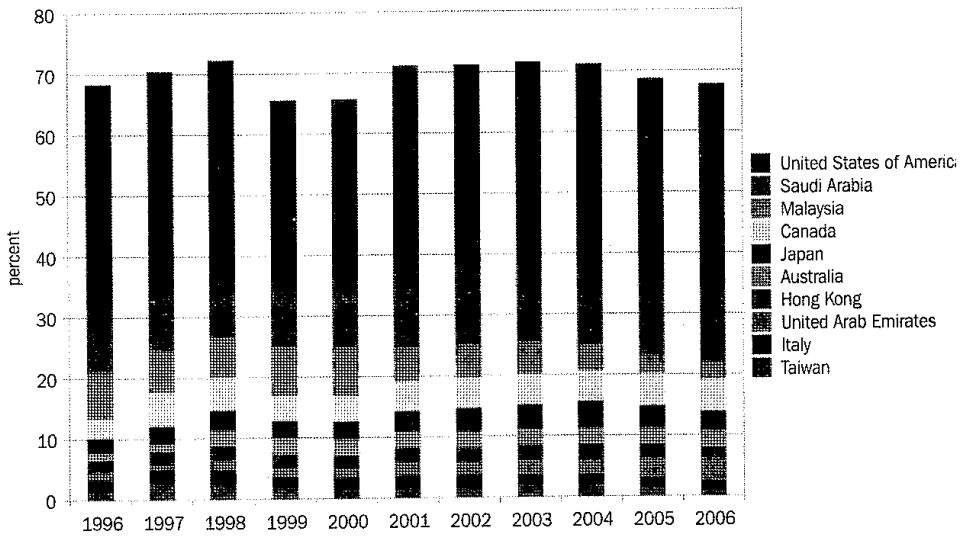
The optimization framework developed in Section III emphasized that since the household supplies labor both to the domestic and international markets, the labor supply decision depends on among other factors, the domestic and international wage rates and the exchange rate. The situation becomes more complicated when the international labor market is heterogeneous, such that markets offer compensation in different currencies. Another implication is that, since the US dollar is typically the currency in which remittances are made, then OFs may choose to retain their incomes in their respective currencies during periods when the peso is strong against the US dollar, and

to remit otherwise. Thus, in this case the decision of when and how much to remit becomes entangled with portfolio allocation decisions, so that the impact of exchange rate movements on behavior becomes more complicated.

Theoretical considerations point out that the relevant exchange rate measure that should enter into the decision-making process should be in real terms, i.e., adjusted for price differentials across economies. However, all the exchange rate indices that have so far been constructed use trade shares as weights for the different currencies. Such indices are more suited to explaining trade flows between countries. Given the rising importance of remittances as a source of foreign exchange for emerging market economies, and given the foregoing theoretical considerations, it may be useful to evaluate the utility of constructing exchange rate indices specifically for the purpose of helping to explain the behavior of remittances.

This section describes our attempt at constructing such indices, and presents estimates that incorporate the index into systems of vector autoregressions. The methodology for computing the OF deployment-based exchange rate index parallels that of the trade-based effective exchange rate indices complied by the BSP, as discussed in Box 1, with OF deployment utilized in place of trade weights. The top 10 countries of destination of OFs from 1996-2005 determined the choice of currencies to be included in the basket as well as the weights attached to them. Figure 6 illustrates the distribution of OFs in the major countries of destination, led by USA, and followed by Saudi Arabia, Malaysia, Canada, Japan, Australia, Hong Kong, United Arab Emirates, Italy and Taiwan. It can be seen that around 70% of OFs, on average, are deployed in these countries. The percent share of each country to total deployment of

Figure 6 **Top 10 Countries of Destination of OFs, 1996-2006 (percent share to total deployment)**



Source: BSP Department of Economic Statistics.

OFs for each year were re-weighted so that individual country weights to be used in the construction of the new exchange rate when summed up will add up to one.

Hence, two deployment-weighted indices were estimated, namely, the nominal OF deployment-based exchange rate (NOER) and the real OF deployment-based index (ROER), or the simply the NOER adjusted for inflation differentials of countries included in the currency basket (using 1980 as the base year, as in the effective exchange rate indices of the BSP).[5] Computationally,

$$NOER_t = \sum_{i=1}^{n} \left\{ \left[\left(\frac{ERPhil_{\$t}}{ERPhil_{\$1980}} \right) \times \left(\frac{ERi_{\$1980}}{ERi_{\$t}} \right) \times 100 \right] \times w_i \right\}$$

$$ROER_t = NOER_t \times \left(\frac{CPI_{Phil}}{\sum_{i=1}^{n} CPI_{i_t} \times w_t} \times 100 \right)$$

where $ERPhil_{\$t}$ is the nominal US$/PhP exchange rate at time t
$ERPhil_{\$t}$ is the nominal US$/PhP exchange rate in 1980
$ERi_{\$1980}$ is the nominal US$/national currency exchange rate of country i in 1980

$CPI_{PHIL\ t}$ is the consumer price index (CPI) level of the Philippines at time t
CPI_{it} is the CPI level of country i at time t
w_{it} is the weight of the currency of country i at time t

The nominal index, NOER, is a weighted arithmetic mean of the changes of the peso relative to each component of a basket of currencies, with the weights being given by annual OF deployment data, while the real index, ROER, adjusts the NOER according to the cost of living differences between the Philippines and each economy whose currency is included in the NOER basket.

Figure 7 compares the behavior of the real index that we have introduced, the ROER index, to that of the real bilateral (₱/$) index where, for comparability, both indices are measured so that an increase in the index reflects an appreciation of the peso. It may be worthy to note that while the Augmented Dickey-Fuller tests show that the ROER and the real bilateral exchange rate indices are both integrated of order one (in levels and logarithms), the Johansen Cointegration Test indicate that the two series are not cointegrated. Although there are some broad similarities in the movements of the two indices, the most striking point is that the peso over the long run has depreciated more against the US dollar than against the OF-

5 Appendix 3 details the weights for the currencies in the NOER and ROER.

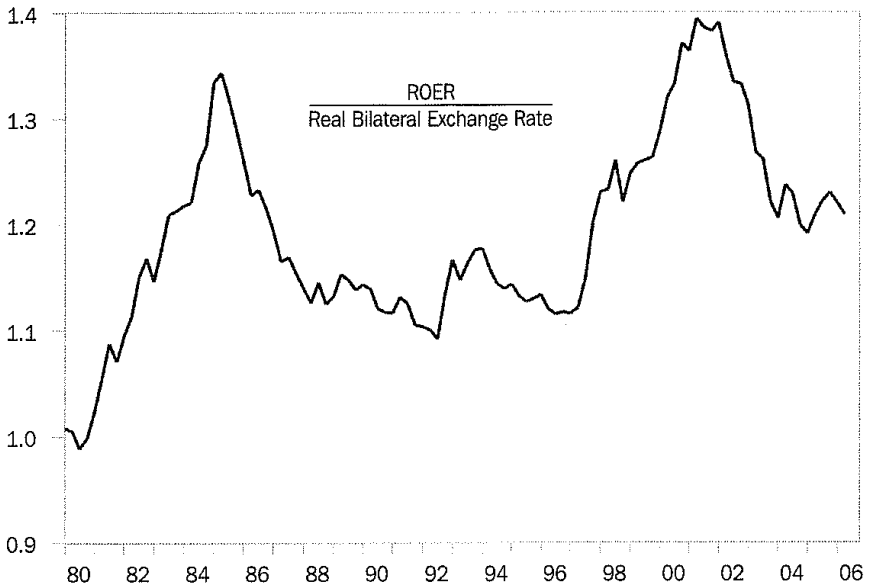
weighted currency basket. The bilateral index is also more volatile: for the period 1980:Q1–2006:Q2, the coefficient of variation for the bilateral index is 0.146, against 0.116 for the ROER index. This is to be expected, since the peso is more likely to share characteristics with the currencies of the OF destination countries than with the US dollar, and there should therefore be greater co-movement between the former group of currencies than between the peso and US dollar. A major implication is that, to the extent that the remuneration of the overseas workers are denominated in the currencies of their destination countries, then OF family incomes are made less sensitive to fluctuations in the peso-dollar rate.

Figure 7 **Real Bilateral (P/$) and Real OF Deployment-Weighted Exchange Rates (ROER)**

Source: Authors' calculations based on the data from the BSP Department of Economic Statistics.

Figure 8 shows the trend, since 1980, of the ratio of the ROER index to that of the real bilateral exchange rate. The ratio can be interpreted as measuring, relative to 1980, the value of the currencies in the ROER basket that can be purchased by one US dollar. It can be seen that, over the long term, the currencies of the OF destination countries have depreciated relative to the dollar. Over the past few years (i.e., since about 2002), however, there has been some reversal, owing to the weakening of the US dollar. It should be noted that this can be one major factor that mitigates any adverse impact of the recent appreciation of the peso on OF family incomes.

Figure 8 **Relative Movements of ROER and Real Bilateral (P/$) Exchange Rate**



Source: Authors' calculations based on the data from the BSP Department of Economic Statistics.

Table 3 **Specification of Estimated Vector Autoregressions**

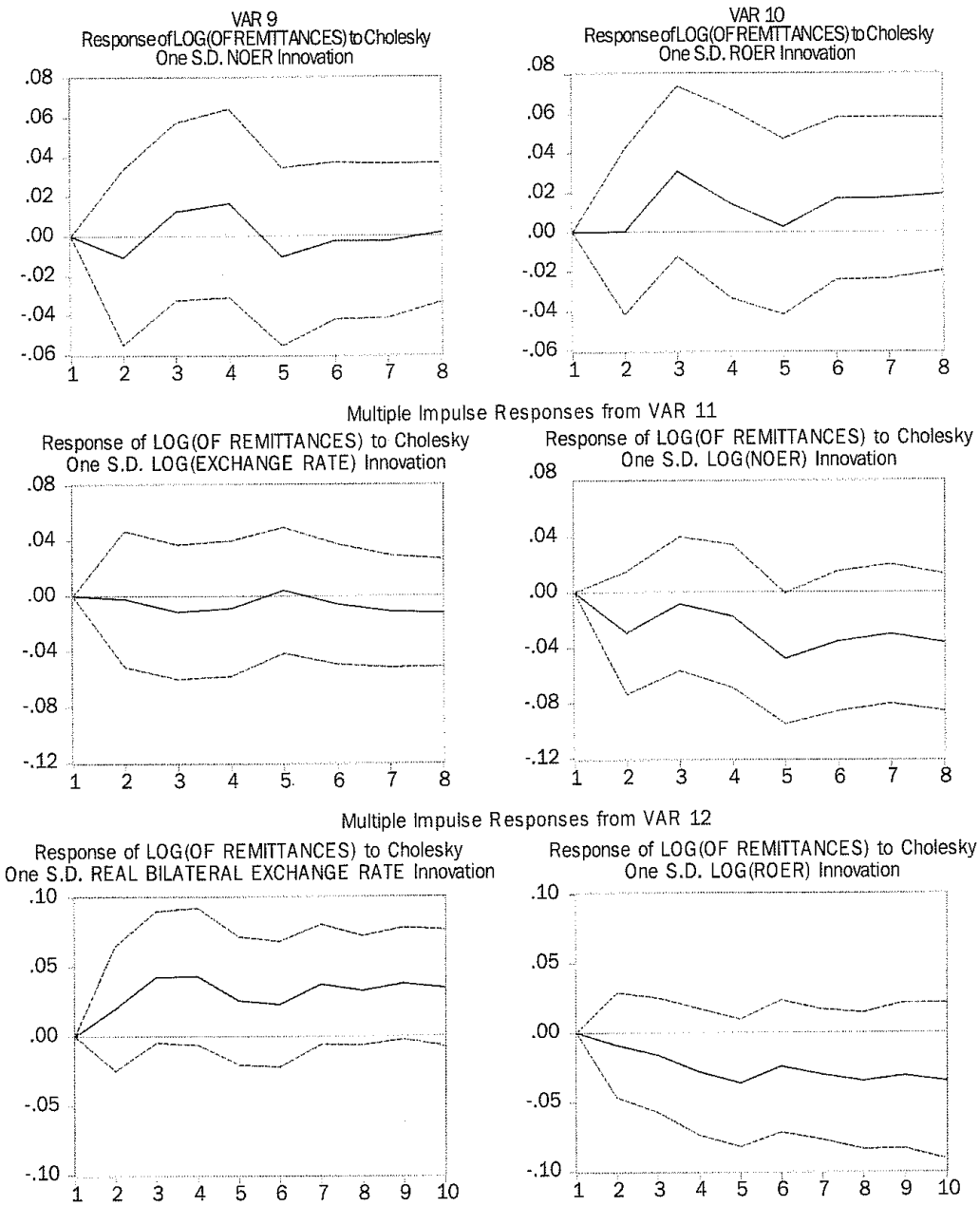
Variable	VAR 9	VAR 10	VAR 11	VAR 12
Log(OF Remittances)	✓	✓	✓	✓
Exchange Rate (P/$)			✓	
Real Exchange Rate ($/P)				✓
NEER Broad				
REER Broad				
NEER Narrow				
REER Narrow				
NEER Major				
REER Major				
NOER	✓		✓	
ROER		✓		✓
Real 91-day T-bill Rate	✓	✓	✓	✓
Log(GDP)	✓	✓	✓	✓
Log(OF Deployment)	✓	✓	✓	✓

Source: Authors' calculations based on the data from the BSP Department of Economic Statistics.

The foregoing considerations argue for inclusion of a deployment-weighted exchange rate index in the vector autoregressions. The specifications of the estimated VARs are summarized in Table 3. VARs 9 and 10 include a single measure of the exchange rate, which is the deployment-weighted measure, in nominal and real terms, while VARs 11 and 12 incorporate two measures simultaneously, one being the nominal or real bilateral ₱/$ rate and the other, the corresponding nominal or real deployment-weighted measure.

The relevant IRFs are presented in Figure 9. Including either the nominal or the real deployment-weighted index as the single exchange rate measure in the VAR fails to produce significant results. In as much as the US dollar still remains the currency of choice for transmission of funds across countries, then failure to include the ₱/$ rate would inadequately account for the exchange rate impact in the VAR system.[6] Hence, it is indeed worthwhile to

Figure 9 **Impulse Responses of OF Remittances from VARs 9-12**



Source: Authors' calculations based on data from the BSP Department of Economic Statistics.

6 The US dollar remains the currency of choice of OFs for sending remittances to the Philippines because it is a common practice of remittance centers in various host countries to course remittances through US correspondent banks.

include both the bilateral and the OF deployment-weighted exchange rates in the estimations, as in VAR 11 and VAR 12, for the nominal and real indices, respectively. The insignificant result for the impact of the nominal peso-dollar rate on remittances in VAR 11 underscores the need to consider differences in the purchasing power of the peso vis-à-vis the US dollar and currencies of major OF host countries. This was addressed accordingly by VAR 12 which yielded the most interesting result, as shown in the lower portion of Figure 9. While an appreciation of the peso relative to the US dollar alone leads to an increase in remittances, the depreciation of the peso relative to the US dollar as well as currencies of major OF destination countries leads to a rise in remittances, indicating the dominance of substitution effect over income effect. This could allow us to peer into the motives of OFs for remitting income earned overseas. As noted in Section II, the dominance of the substitution effect is more in line with investment than with altruistic motives as the increase in remittances brought forth by the depreciation can be seen as a way to take advantage of investment opportunities in the Philippines.

7. Conclusion and Policy Implications

To answer the question posed by the title of this paper, we showed that an OF deployment-weighted exchange rate index proved to be a significant determinant of remittances, both at the empirical and theoretical levels. A depreciation of the peso vis-à-vis a basket of currencies of major OF destination countries prompts an increase in remittances. This indicates that OFs are driven by investment-related rather than altruistic motives.

The findings of this paper carry important implications. On the significance of an OF deployment-weighted exchange rate index as a determinant of OF remittances, it may be worthwhile to evaluate the utility of constructing exchange rate indices specifically for the purpose of helping to explain the behavior of remittances. On the predominance of investment over altruistic motives of OFs, it underscores the need to develop investment programs that could induce OFs to channel their remittances toward productive investments. To enhance the impact of remittances on savings, investment, and thus economic growth, it is vital to strengthen the incentives to "bank the unbanked". The BSP has adopted measures to encourage overseas Filipinos to remit through the financial system. These measures are anchored on four principles such as enhancing transparency and promoting competition in the remittance market; improving the country's payment and settlement systems and the access to financial services; encouraging OFs and their families to increase savings and investments; and cultivating financial literacy among

OFs and their families.[7] Finally, it is worthwhile to note that OFs will react in the same way as local residents and foreign investors to poor macroeconomic policies–they will reduce local exposure if these policies persist. Thus, authorities need to preserve macroeconomic stability so that remittances will continue to have a positive impact on the economy.

As with previous studies that analyzed the determinants of remittances, this paper calls on future studies to treat remittances as another macroeconomic variable that is endogenous to other home country variables and not as an exogenous flow of money from abroad. Given the significant impact of exchange rate changes to remittances and the profit-driven motives of OFs, as this paper has shown, remittance flows can affect saving and investment behavior and thus future growth. The magnitude of these links depends on many direct and indirect effects and are largely determined by the structural features of the economy and the relevant elasticity values. Hence, tracing these effects requires a general equilibrium model, such as a dynamic stochastic general equilibrium model (DSGE), where endogenous labor supply decisions of OFs are explicitly incorporated and enough sectoral detail exists to allow the transmission of macroeconomic variables affecting remittances to economic growth.[8]

7 In particular, the measures adopted by the BSP to encourage OFs to remit through formal channels include, among others, the issuance of Circular No. 534 (disclosure of remittance charges and other relevant information) and Circular No. 564 (requirement of valid identification for financial transactions); inclusion of OFW portal in the BSP website; grant of foreign currency deposit unit (FCDU) license to rural banks and cooperatives; interconnection of major automated teller machine (ATM) networks in the country; approval of alternative modes of remittances; and conduct of Financial Literacy Campaign (FLC) among OFs and their beneficiaries.

8 To this end, initial steps have been undertaken in Dakila and Dakila (2006) which examined the impact of remittances in a multi-region computable general equilibrium model.

Appendix 1 Studies on Impact of Exchange Rate on Remittances			
Author/s	Coverage		Impact of Depreciation on Remittances
	Time Period	Countries	
Faini (November 1993)	1971-1989	Morocco Portugal Tunisia Turkey Yugoslavia Greece Italy Portugal Spain Turkey Yugoslavia	+
Loser, et al. (2003)	unspecified	Colombia, The Dominican Republic El Salvador Mexico	Inconclusive (positive only for a subsample of countries covered by the study that included Columbia and El Salvador)
Higgins, et al. (2004)	1970-1997	Bolivia Columbia Dominican Republic El Salvador Guatamela Honduras Jamica Mexico Peru	Not significant
Yang (June 2006)	July 1997 – October 1998	Philippines	+
Faini (October 2006)	1990, 2000	38 countries in the Caribbean, Americas, Africa and Asia	Not significant
Lueth and Ruiz-Arranz (2006)	1980-2004	11 countries in Asia, Europe and Middle East	-
Vargas-Silva (March 2007)	January 1995-February 2006	Mexico	Not significant

APPENDIX 2: Selected Studies on the Cyclicality of Remittances			
Author/s	Coverage		Cyclicality of Remittances
	Time Period	Countries	
Chami, et al. (2003, 2005)	1970-1998	113 countries including the Philippines	Countercyclical
BSP-DER (2003)	1970-2003	Philippines	Inconclusive
Burgess and Haksar (2005)	1985-2002	Philippines	Inconclusive
Tuaño-Amador, et al. (forthcoming)	1989-2006	Philippines	Procyclical

Appendix 3 **Currency Weightings in the NOER and ROER[1]**

Currency	1996[2]	1997	1998	1999	2000	2001	2002	2003	2004	2005
US dollar	0.55	0.53	0.52	0.45	0.45	0.49	0.49	0.48	0.49	0.50
Saudi riyal	0.15	0.13	0.13	0.18	0.18	0.18	0.18	0.18	0.18	0.19
Malaysian ringgit	0.12	0.10	0.09	0.13	0.13	0.08	0.08	0.08	0.06	0.05
Canadian dollar	0.05	0.08	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.08
Japanese yen	0.04	0.04	0.04	0.04	0.04	0.05	0.06	0.06	0.06	0.06
Australian dollar	0.02	0.02	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Hong Kong dollar	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.04	0.04
Italian lire[3]	0.02	0.04	0.04	0.03	0.03	0.02	0.02	0.03	0.03	0.02
New Taiwan dollar	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.02	0.03	0.02
Total	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

1 United Arab Emirates dirham was excluded in the basket of currencies due to lack of available continuous data series. The remaining nine currencies of top OF destination countries were re-weighted so that the weights add up to one.

2 The currency weightings for 1996 were applied to the earlier period (1980-1995) due to unavailability of OF deployment data for said period.

3 With the introduction of the euro in 1999, a conversion factor obtained from the European Central Bank website <http://www.ecb.int/bc/intro/html/index.en.html#fix> was used to generate a continuous data series for the Italian lire until 2005.

References

Bangko Sentral ng Pilipinas (2003), Department of Economic Research (DER) Validation Exercises of IMF Study "Are Immigrant Flows a Source of Capital for Development?" by R. Chami, C. Fullemkamp and S. Jahjah (unpublished).

Bussolo, M. and Medvedev D. "Do Remittances Have a Flip Side? A General Equilibrium Analysis of Remittances, Labor Supply and Policy Options for Jamaica".

Burgess, R. and Haksar V. (2005), "Migration and Foreign Remittances in the Philippines." *IMF Working Paper* No. 05/111.

Chami, R., Fullenkamp C., and Jahjah S. (2003), "Are Immigrant Remittance Flows a Source of Capital for Development?" *IMF Working Paper* No. 03/189. Available at SSRN: http://ssrn.com/abstract=463002 or DOI: 10.2139/ssrn.463002.

Coughlin, C. and Pollard, P. (1996), "A Question of Measurement: Is the Dollar Rising or Falling?" *Review*, July/August, Federal Reserve Bank of St. Louis.

Dakila, C. and Dakila, F. (2006), "Modeling the Impact of Overseas Filipino Workers on the Philippine Economy: An Inter-Regional and Economy-Wide Approach." *BSP Working Paper Series No. 2006-002.*

Faini, R. (1994), "Workers' Remittances and the Real Exchange Rate". *Journal of Population Economics*, 7:235-245.

_____ (2006), "Migration and Remittances: The Impact on the Countries of Origin." Paper presented at EUDN-AFD conference on Migration and Development on 8 November 2006 in Paris, France.

Hagen-Zanker, J. and Siegel M. (2007), "The Determinants of Remittances: A Review of the Literature." *Maastricht Graduate School of Governance Working Paper* No. 3.

Higgins, M., Hysenbegasi, A., and Pozo, S. (2004), "Exchange Rate Uncertainty and Workers' Remittances." *Applied Financial Economics,* 14:403-411.

Loser, C., Lockwood, C., Minson, A., and Balcazar, L. The Macroeconomic Impact of Remittances in Latin America: Dutch Disease or Latin Cure?

Lucas, R. E. B. and Stark, O. (1985), "Motivations to Remit: Evidence from Botswana." *The Journal of Political Economy, 93(5), 901-918.*

Lueth, E. and Ruiz-Arranz M. (2006), "A Gravity Model of Workers' Remittances." *IMF Working Paper* No. 06/290.

Schiopu, I. and Siegfried, N. (2006), "Determinants of Workers' Remittances: Evidence from the European Neighboring Region." *Working Paper Series* No. 688, European Central Bank.

Tuaño-Amador, C., Claveria, R., Delloro, V., and Co, F. "Philippine Overseas Workers and Migrants' Remittances: The Dutch Disease Phenomenon and the Cyclicality Issue." *Bangko Sentral Review.* 2007 Volume IX, No. 1.

Yang, D. (2006), "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks." *University of Michigan Working Paper.*

Vargas-Silva, C. (2007), "The Tale of Three Amigos: Remittances, Exchange Rates and Money Demand in Mexico". Available at SSRN: http://ssrn.com/abstract=956107.

Population Dynamics and Household Savings: Evidence from the Philippines

Dennis S. Mapa[1] and Lisa Grace S. Bersales[2]

Abstract

The economic growth implications due to changes in the nation's age structure have been substantial. In the course of the demographic transition, countries experience an increasing share of the working age population relative to the total population and this creates favorable effects on economic growth. The changing age structure also influences household saving rate. Household saving rate in the Philippines is one of the lowest in East Asia. This paper looks at the impact of the slow demographic transition in the Philippines on its aggregate household saving rate using panel data from the Family Income and Expenditure Survey (1985 to 2003). The econometric model is based on the augmented life cycle model and the results suggest that the country's population dynamics play an important role in its household saving rate. The Philippines' rapid population growth formed a big bulge at the lower portion of the age pyramid that resulted in a higher percentage of young dependents. The data suggests that the country's high population growth resulted in low household saving rate and consequently, low economic growth. The study also shows that remittances from migrant workers are a major source of aggregate household savings.

Keywords: Demographic transition, household saving rate, augmented life cycle model, population dynamics.

1. Introduction

The economic growth implications due to the changes in the nation's age structure, resulting from the demographic transition, have been substantial and are of interest in research in recent years. Studies, notably that of Bloom and Canning (2001) and Bloom and Williamson (1998), show that demographic factors have strong and significant effects on economic growth. Demographic transition is described as "a change from a situation of high fertility and high mortality to one of low fertility and low mortality." A country that enters into a demographic transition experiences sizable changes in the age distribution

1 Associate Professor and Director for Research, School of Statistics, University of the Philippines Diliman. Email address: cdsmapa@yahoo.com.

2 Professor, School of Statistics, University of the Philippines Diliman.

of the population. Demographic transition has three phases and each phase has a different impact on the economy. Phase one is triggered by an initial decline in infant mortality but fertility remains high resulting in the swelling of the youth dependency group (0 to 14 years) as well as the demand for basic education and primary health care. This phase poses a big challenge to the economy as it may hinder economic growth. It should be noted that the Philippines has been stuck in the first phase of the demographic transition for the last 40 years. In the second phase of the transition, these "baby boomers" enter the adult labor market (some 20 years later) and if the market is able to absorb them, they can accelerate the phase of economic growth. This is the phase when the proportion of working-age population is highest and the age dependency ratio or the ratio of young dependents (0 to 14 years) and elderly (65 years and above) over the working age (15 to 64 years) is lowest. Countries that are currently in the second phase of the demographic transition are Singapore, South Korea, Taiwan, and Thailand. The third and last phase of the transition occurs when the elderly cohort (65 years and above) swells relative to the total population. An example of a country currently at the third phase of the demographic transition is Japan.

In the course of the demographic transition, countries experience an increasing share of the working age population relative to the total population and this creates favorable effects on the per capita income. Mason and Lee (2006) refer to this effect of the demographic transition to income growth as the "first dividend." The changing age structure interacts with the life cycle of production and consumption. Young (ages 0 to 14 years) and elderly (ages 65 years and above) members of the population produce less than they consume while the working age adults (ages 15 to 64 years) produce more than they consume. Hence, countries having a population structure with heavy concentration at the working–age group have the advantage of producing high levels of per capita income, all things being the same. However, the impact of the first dividend is conditioned on policies related to the labor market; that is, how wages and labor force participation rates react to the rapid increase in the working-age population. Cross-country and intra-country econometric analyses (Mapa and Balisacan 2004; Mapa, Balisacan, and Briones 2006) have shown that the Philippines has not benefited from the so-called demographic dividend that is a major contributor to the economic success experienced by East Asian countries from the 1960s to 1990s.

Mason (2007) discusses another form of dividend resulting from the changing age structure of the nation's population and refers to it as the second demographic dividend. The second dividend results from the society's response to the prospect of an aging population, an outcome as the nation's age structure enters into the third phase of the demographic transition. A challenge faced by societies (and governments) when there is a substantial percentage of the elderly population is how to support their consumption,

given a reduction in their income. There are several approaches to this problem. These include: (a) relying on public (or familial) transfer systems and (b) increasing saving rates and accumulating greater physical wealth or capital. Individuals accumulate savings in their working age years which serve as buffer during their retirement years. While accumulation of capital can be used to deal with the life cycle deficit in the older years, this capital also influences economic growth. As Mason points out, when society increases its saving rate, it also experiences more rapid economic growth – resulting in the second demographic dividend.

The link between savings, population, and economic growth is inherent in the neoclassical growth model of Solow (1956) and Swan (1956). In the model, output per worker is determined only by two variables: capital per worker and the level of technology. The production function is described by, $Y_t = F[K_t, A_t L_t]$ where K_t is the physical capital, L_t is labor and A_t is the labor-augmenting technical progress. The constant return to scale assumption means that the production function can be expressed in intensive form (in per worker or per capita form), $\frac{Y}{AL} = F\left[\frac{K}{AL}, 1\right]$ or y = f(k), where k is the capital per worker and y is the output per worker. Labor and knowledge grow at constant exogenous rates, $\dot{L}_t = nL_t$ and $\dot{A}_t = gA_t$. Capital accumulation is determined by savings, $S_t = sY_t$, where s is fixed. The condition for a change in capital stock is that, $\dot{K}_t = sY_t - \delta K_t$, where δ is the depreciation rate. The fundamental result of the Solow-Swan model is,

$$\dot{k}_t = sy_t - (n + \delta)k ,$$

where the first term on the right side of the equation is the amount of new capital being provided each period by the average worker. The second term is indicative of the effective depreciation rate for the capital-labor ratio. If the saving rate is zero, then the capital per person will decline partly because of the depreciation of capital (at rate δ) and partly due to an increase in the number of persons (at the rate n). If savings exceed the necessary amount to equip new workers, then the capital-labor ratio increases and capital-deepening occurs (Lee, Mason, and Miller 2001).

The model yields three important implications that relate savings, population, and economic growth. First, an increase in the saving rate or a decrease in the population growth yields a higher equilibrium output per worker. Second, an increase in the saving rate or a decrease in the population growth produces a transitory increase in the growth rate of output per worker. And lastly, at the steady state neither saving rate nor population growth rate has an impact on the rate of growth of output per worker.

The two sources of capital-deepening in the Solow-Swan model that experienced high economic growth rates from the 1960s up to the 1990s (Lee et al. 2001) are: (a) rise in saving rate and (b) decline in population growth have been observed in the East Asian economies (Singapore, South Korea, Taiwan, and Thailand). Using the cross-country data, Mapa and Balisacan (2004) show that the average gross domestic saving ratios during the period 1975 to 2000 for these countries are 28% for Thailand, 32% for South Korea, and 44% for Singapore. The gross domestic saving (GDS) rate of the Philippines for the same period is only 22%.

This paper looks at the link between population dynamics or the changing age structure and the saving rate in the Philippines using aggregate (regional) household panel data generated from the Family Income and Expenditure Survey (FIES) from 1985 to 2003. The research is motivated by the fact that unlike its neighbors, the Philippines has failed to benefit from the second demographic dividend, where high saving rates lead to an even higher economic growth. The study makes use of econometric models to explain the connection between the population dynamics and household saving rate. The econometric model is based on the augmented life cycle model.

The remainder of the paper is organized as follows: Section 2 presents empirical studies and simulated results on the population dynamics-saving rate-economic growth nexus in the countries that experienced demographic transition. Section 3 presents a profile of household savings in the Philippines using data from the 1985, 1988, 1991, 1994, 1997, 2000, and 2003 FIES. The theoretical framework of the econometric model for saving rate using panel data is discussed in Section 4. Section 5 presents the empirical results and Section 6 states the conclusion.

2. Population Growth-Saving Rate-Economic Growth: Country Studies

The relationship between population dynamics and saving rate is integrated in the life cycle model of consumption. Modigliani (1986) asserts "the self-evident proposition that the representative consumer will choose to consume at a reasonably stable rate, close to his anticipated average life consumption." The life cycle model predicts that both demographic variables and productivity growth will generate saving–the young save while the elderly dissave; and if it is assumed that the population is stationary with the income of the young the same as the income of the old, then saving and dissaving will be equal and opposite (Deaton 1992). However, a different picture emerges if a country enters into a demographic transition. During the first phase of the demographic transition the young dependent population (ages 0 to 14

years) is growing faster relative to the working-age population (ages 15 to 64 years) resulting in higher household consumption, which in turn diminishes the rate of savings (Coale and Hoover 1956). During the second phase of the demographic transition, working-age population is growing faster relative to the young dependent population resulting in higher saving rates.

A series of empirical studies based on cross-country aggregate-level panel data show that demographic factors have a strong and statistically significant effect on aggregate savings (Bloom, Canning, and Graham 2003; Deaton and Paxson 2000, and Kelly and Schmidt 2007). Country-specific studies also validate the results of the cross-country studies. One such example is the series of studies that used similar data sets for Taiwan. Williamson and Higgins (2001) and Kelly and Schimdt (1996) have shown that changes in the age structure over the demographic transition brought about an increase in the gross national saving rate of Taiwan by 25 percentage points (Kelly and Schmidt 1996) to 45 percentage points (Williamson and Higgins 2001). However, using the same data for their analysis, Lee, et al. (2001) estimated only about 14.5 percentage points increase in saving rate due to age structure. Lee, et al. (2001) suggest that the seemingly high impact of demographic factors on household saving rate, particularly that of Williamson and Higgins, is due to the fact that the authors essentially relate all the increase in the household saving rate to the demographic transition. Lee, et al. argues that there are short-term fluctuations in the household saving rate that are non-demographic related. On the other hand, Deaton and Paxson's (2000) estimate of the increase in household saving rate in Taiwan induced by the demographic transition is around 6.5 percentage points only. Deaton and Paxson's analysis indicates that most of the increases in the household saving rate experienced in Taiwan are due to non-demographic factors, which they ascribe to changes in cohorts and time effects (Lee, et al. 2001). While there is a need to reach a consensus on the impact of demographic factors on household saving, empirical results suggest that the estimated rise in household saving rate accounted for by the demographic transition is economically significant.

Lee, et al. (2001) also made use of simulations to compare the impact of saving rate (due to the demographic factors) on per capita income in Latin America and Taiwan. The comparison is relevant because while the fertility decline in Latin America begun in about the same year as in Taiwan, the transition to replacement fertility took 60 years in Latin America compared to 30 years in Taiwan. The simulation results show that the Latin American scenario resulted in the same per capita income level as that in Taiwan but several decades later. The authors concluded that for countries that experience rapid demographic transitions, the saving rates remain high for

several decades corresponding to a rapid increase in the per capita income growth.

Another simulation analysis of Mason (2001) using household data from Taiwan shows that the high rate of savings and investment resulting from the demographic transition accounts for 18% of the increase in the output per worker during the period 1960 to 1990, supporting the notion of the second demographic dividend (higher saving rate due to the demographic transition). Both the first demographic dividend (from the gap between population and labor force growth or the translation component), and the second demographic dividend account for about 27.7% of the increase in the per capita output in Taiwan. In Mason's simulation, using the middle-of-the-road estimate of the saving effects, the second demographic dividend (18%) is even higher than the first demographic dividend (9.7%).

Mason (2007) uses simulations to estimate the wealth accumulation of three countries, India, Japan, and the United States during the course of the demographic transition. These countries have experienced demographic transitions in various periods: India from 1975 to 2000, Japan from 1950 to 1980, and the United States during the period 1850 to 1940. Noting that wealth is accumulated during the working age years, Mason introduces the concept of total life cycle wealth or the wealth held by all individuals older than the age at which the accumulation process begins which the author estimates to be at 50 years. Mason assumes that surplus at younger working age is transferred to children for their consumption and thus wealth accumulation starts only when the person reaches the age 50 when the transfer to children stops. (Note that in the case of the Philippines, the intergenerational support stops at an age higher than 50.) Data from the 2000 Census of Population and Housing (CPH) show that more than half (57%) of the population 60 years and above are household heads. Moreover, about 72% supports at least one family member (other than his/her spouse). There are even household heads aged 60 years and above (12%) that support more than eight family members, mostly children and grandchildren. Mason's simulations show that by 1950, the United States life cycle wealth relative to total income is four times that of India and about twice that of Japan. During this period, Japan's simulated life cycle wealth begins to increase very rapidly and that by 1975, the ratio of life cycle wealth to output is 6.0 in Japan and only 3.6 in the United States. In the case of India, the simulated life cycle wealth increased very slowly until recently. The ratio of life cycle wealth to income did not reach 1.0 until 1985 (Japan reached this milestone in 1940 and the United States in 1905).

In the same paper, Mason also provides estimates of the effects of the two demographic dividends on the average per capita GDP growth of the three countries. The first dividend is due to the translation effect on income growth (faster rate of working-age population compared to the total population) and

the second dividend is due to the effect of capital accumulation on income growth. For Japan, during the period 1950 to 1980, the simulations show that of the average yearly per capita GDP growth rate of 6.23%, the first dividend accounts for 0.63% and the second dividend accounts for 1.72%. The two dividends due to demographic transition account for 37.7% of the yearly average per capita growth rate of Japan. Using the data from the United States from 1850 to 1940, the simulations show that the first dividend and the second dividend account for 0.27% and 0.62%, respectively, of the average yearly per capita GDP growth rate of 1.55% (57.6% of the actual). For India during the period 1975 to 2000, the simulations show that the first and second demographic dividends account for 0.34% and 1.02%, respectively, of the yearly average per capita GDP growth rate of 2.98%.

In another cross-country analysis, using data from 112 developing economies and 22 industrial economies for the period 1965 to 1995, the World Bank (1999) shows that an increase in the share of the young dependents in the population tends to reduce private savings. The study pointed out that an increase in the young age dependency ratio of 3.5 percentage points leads to about a one percentage point decline in private savings. The study concluded that "developing countries undergoing a demographic transition, in which the working-age population is a large and growing share of the population, may witness a transitory increase in their saving rates."

In the Philippines, Rodriguez and Meyer (1988) examined the saving behavior of 1000 rural households using data gathered by the Agricultural Credit Policy Council (ACPC) of the country in 1987. The authors found that factors such as income, household size and education of the household head, among others, played significant and positive roles in raising savings among rural households. Bautista and Lamberte (1990) analyzed the saving behavior of rural and urban households in the Philippines using data of 17,495 households from the 1985 FIES, and found that the marginal propensity to save for households in Metro Manila is lower than that for households in any other region (except Region 2). The study also showed that at a given income level, rural households generally save more than urban households. Moreover, the marginal saving rate of rural households increases more rapidly as they move up from low- and middle-income groups to the high-income group compared to the urban households. Both studies, however, did not include demographic variables in their econometric models.

3. Household Saving Profile in the Philippines (1985 to 2003)

The saving patterns of households obtained using the relevant data from the FIES for the years 1985, 1988, 1991, 1994, 1997, 2000, and 2003. The number of households surveyed is presented in Table 1. The regional classification

Table 1 **Sample Households by FIES Year**

FIES Year	Number of Sample Households
1985	16,971
1988	18,922
1991	24,789
1994	24,797
1997	39,520
2000	39,615
2003	42,094

Source: National Statistics Office, Philippines: Family Income and Expenditure Survey data, various years.

Table 2 **Regions in the 1988 Family Income and Expenditure Survey (FIES)**

Region 1	Ilocos Region
Region 2	Cagayan Valley
Region 3	Central Luzon
Region 4	Southern Tagalog
Region 5	Bicol Region
Region 6	Western Visaya
Region 7	Central Visayas
Region 8	Eastern Visayas
Region 9	Western Mindanao
Region 10	Northern Mindanao
Region 11	Southern Mindanao
Region 12	Central Mindanao
Region 13	National Capital Region (NCR)
Region 14	Cordillera Autonomous Region (CAR)

Source: National Statistics Office, Philippines: Family Income and Expenditure Survey data, various years.

uses the fourteen regions (defined in the 1988 FIES) and is presented in Table 2.

It should be noted that the data set used in the econometric analysis includes only 14 regions, instead of the current 17 regions. The geographical boundaries of the regions (defined in 1988) were kept constant throughout the period 1985 to 2003 for consistency in the definition of the regional units.

Households are also categorized by per capita income deciles at the national and regional levels in the case of regional profiles. Accordingly, households are first ranked nationwide according to their per capita income and then categorized into the income deciles. This classification was maintained even in the regional profiling of saving behavior. Household savings rather than aggregate savings are the focus of this paper. This treatment allows the analysis of saving behavior under the life cycle model and for the different

per capita income deciles as well as by region. The FIES data supports these objectives. Aggregate savings computed from macroeconomic data such as Gross Domestic Product (GDP), gross regional domestic product (GRDP), and the flow of funds data cannot provide micro-level data needed for such an analysis. This is the same reason for the use of household savings rather than aggregate savings in Attanasio and Szekely's (2001) study.

The following operational definitions of household savings and saving rate, respectively, are used in the study:

1. Saving = (Aggregated income of all households – Aggregated expenditure of all households)

2. Saving rate = $\frac{\text{Saving}}{\text{Aggregated income of all households}} \times 100$

Total family income and total family expenditure of the FIES are used in computing savings and saving rate. Total family income includes total wages and salaries, pensions, dividend from investments, interests, rentals, cash receipts/gifts/support from domestic and international sources, net share of crops, and income from family sustenance activities as well as receipts from other sources not elsewhere classified.

The definition of savings used in this paper has been fully discussed (e.g., Attanasio and Szekely 2001). However, as in other studies, this definition of savings has been used due to the limitations in the collection of data of household consumption of durables. Another operational definition of savings is to take out from expenditures the amount of durables and other items consumed which may be viewed as household investment. In the FIES questionnaire, however, no disaggregation of such items is done.

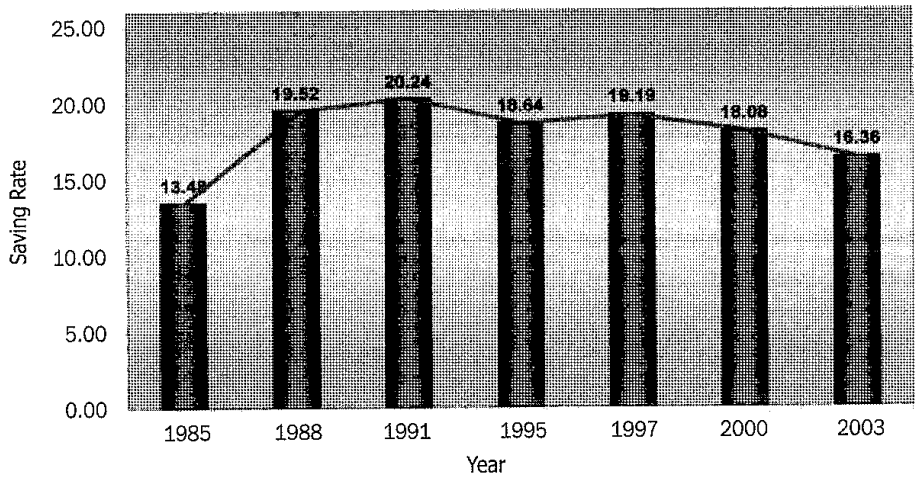
In order to compare savings across FIES years, the data is deflated using the consumer price index (CPI) (1997=100). The Cost of Living Index (reference=NCR) developed by the Asia Pacific Policy Center (APPC) was used to adjust savings to compare it across the regions.

A. Saving Rate

Based on 2003 FIES, household saving rate nationwide is 16%. In 1997, it was 19%. This is lower than the recorded household saving rate of Thailand of 30% and Taiwan of 49% in 1996 (Attanasio and Szekely 2001). Recent years' saving rates are higher than the 1985 level. However, from 1997, a downward trend is noted, as shown in Figure 1.

The downward trend in household saving is generally reflected by the saving rates across the different national per capita income deciles, except for

Figure 1 **National Saving Rate, FIES Years 1985 to 2003**



Source: National Statistics Office, Philippines: Family Income and Expenditure Survey data, various years.

Table 3 **Household Saving Rate by FIES Year and National Per Capita Income Decile**

FIES Year	National Per Capita Income Decile									
	1	2	3	4	5	6	7	8	9	10
1985	-15.02	-2.95	-0.58	2.78	4.69	7.87	8.53	11.34	14.48	26.65
1988	-10.45	0.13	4.70	8.28	10.53	12.65	15.17	17.20	20.94	33.53
1991	-10.97	2.32	5.96	10.44	10.97	13.42	17.03	17.18	21.47	32.34
1994	-6.77	2.89	5.91	9.67	11.53	12.56	15.33	17.91	20.96	28.97
1997	-11.60	-1.20	2.25	6.49	8.58	11.26	14.58	18.13	20.96	30.85
2000	-8.29	-1.58	2.26	5.84	8.77	11.69	13.68	17.63	20.45	28.62
2003	-9.38	-1.64	1.66	4.51	6.95	9.66	11.78	14.52	17.46	29.42

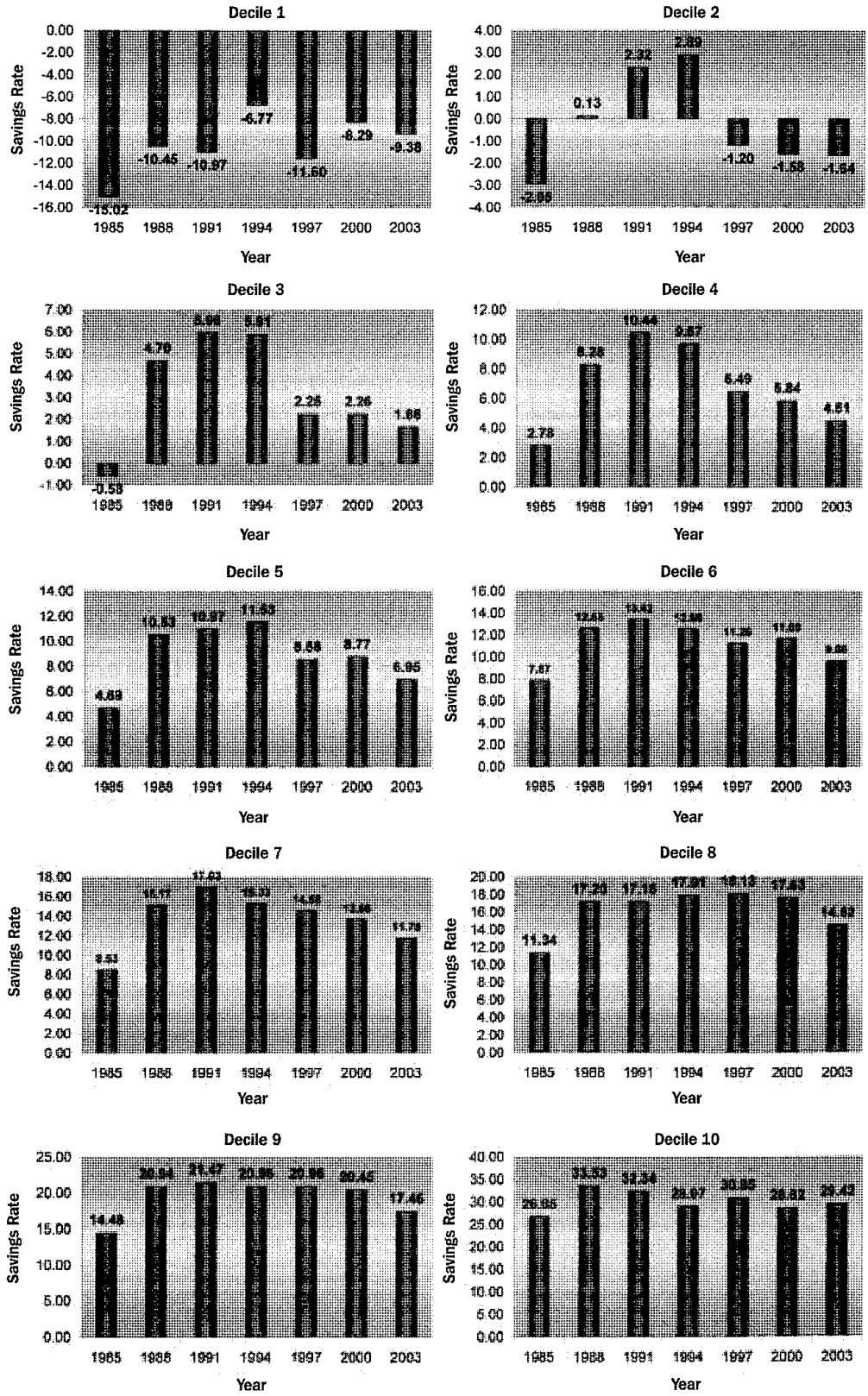
Source: National Statistics Office, Philippines: Family Income and Expenditure Survey data, various years

the highest income decile, whose saving rate increased in 2003 from 2000. Table 3 further shows that the bottom 20% reflects dissaving. It is also noted from this table that the income decile with the highest saving rate is 1.7 times that of the second-ranking income decile and is more than twice that of the other income deciles. Table 3 and Figure 2 present the saving rates and visual illustration of the saving patterns, respectively.

B. Life Cycle Profiles

The FIES data collected by the National Statistics Office (NSO) is a pooled data set of households where a sample of predetermined number of households is collected every three years. The FIES series does not collect genuine panel

Figure 2 **Household Saving Rate by FIES Year for Different Income Deciles**



Source: National Statistics Office, Philippines: Family Income and Expenditure Survey data, various years.

data of a household, that is, the FIES is not a collection of data from the same household through time. The lack of panel data creates a problem in the analysis of household savings since savings is a dynamic phenomenon. To assuage this problem, the authors made use of the synthetic cohort techniques pioneered by Browning, Deaton, and Irish (1985) and adopted by Attanasio and Szekely (2001) in their analysis of household saving in East Asia and Latin America.

The basic idea of the synthetic cohort analysis is to follow the average behavior of a group of households, rather than the individual household. The group membership is assumed to be fixed over time. This strategy allows the author to study the dynamic behavior of the average household saving rates through time. In the study of the aggregate household savings, the synthetic cohort analysis is used, by grouping the different households, according to the age group of the household head, from different FIES years (1985 to 2003) and the average behavior of these groups are assumed representative of cohort behaviors through time. This type of analysis is, however, not immune to problems as pointed out by Attanasio and Szekely (2001), particularly the endogeneity of family formation and dissolution, differential mortality and migration rates across the different socio-economic groups.

Table 4 shows the national saving rate by age group. It is noted that those aged 85 and above are the ones with the higher saving rates. The regions with the highest saving rates in these age groups are Ilocos Region,

Table 4 **Nationwide Saving Rate, by Age Group: FIES Years 1985-2003**

Age Group	Average Saving Rate for all Regions	Highest Regional Saving Rate	Region with Highest Saving Rate
15-19	11.1	30.1	Eastern Visayas
20-24	14.5	20.8	Ilocos Region
25-29	14.0	18.3	Cordillera Autonomous (CAR)
30-34	15.5	19.6	National Capital (NCR)
35-39	14.9	19.2	Cordillera Autonomous (CAR)
40-44	15.7	19.4	Eastern Visayas
45-49	17.4	20.7	Cagayan Valley
50-54	20.3	32.8	Southern Mindanao
55-59	21.8	26.0	National Capital (NCR)
60-64	22.0	32.2	Northern Mindanao
65-69	18.0	26.8	Cordillera Autonomous (CAR)
70-74	24.6	32.2	National Capital (NCR)
75-79	18.9	27.5	Central Mindanao
80-84	19.0	28.9	Western Mindanao
85-89	21.3	30.2	Ilocos Region
90-94	22.2	32.3	Western Visayas
95-99	28.7	37.2	Southern Luzon

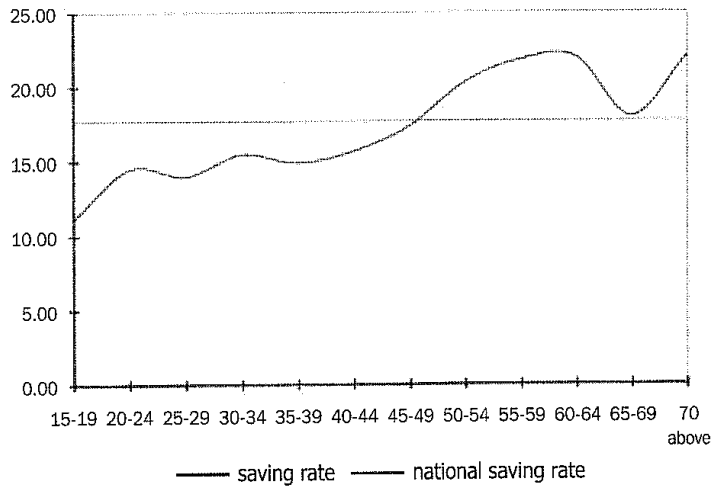
Notes: Highest Saving Rate = 28.7%; Lowest Saving Rate = 11.1%;
Standard Deviation = 4.4; C.V. = 23.5

Source: National Statistics Office, Philippines: Family Income and Expenditure Survey data, various years.

Western Visayas, and Southern Luzon. Those with ages 70 to 74 follow with a nationwide saving rate of 24.6%. The region having the highest saving rate for this age group is NCR. The next age groups with saving rate 20% or more are those with ages 50 to 64. The regions with the highest saving rates for these age groups are Southern Mindanao, NCR, and Northern Mindanao. In younger age groups, the regions with highest saving rates are Eastern Visayas, Ilocos Region, CAR, and NCR.

The life cycle model is validated by the FIES data as shown in Figure 3, except for the age group 70 years and above, that failed to exhibit the expected dissaving. The information in this figure is derived from the data for the FIES years 1997 to 2003. It should be highlighted that the saving rate peaked at the age group 50 to 64 years, which is relatively late compared to Taiwan and Thailand. In both these countries, savings start its peak at 40 to 44 years. The 2000 Census of Population and Housing revealed that 72% of the Philippine population belonged to ages 34 years and below, while only

Figure 3 **Household Saving Rate by Age Group of Head of Household (1997 to 2003)**

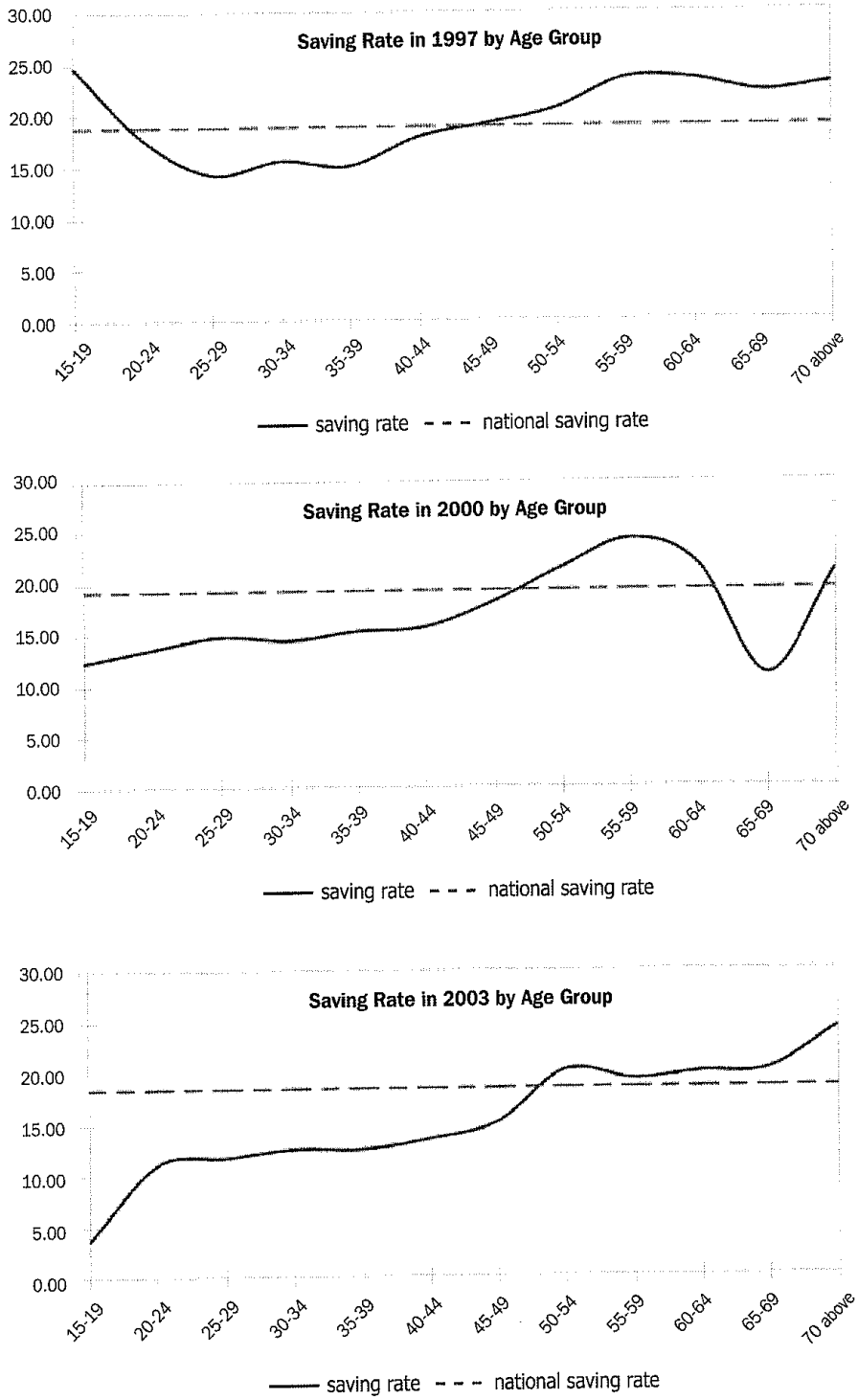


Source: National Statistics Office, Philippines: Family Income and Expenditure Survey data, various years.

8% belonged to the age group, 50 to 64 years. Thus, the household saving rate was substantial at least for the years covered by this analysis.

The upward blip of saving rate for age group 70 years and above appears to hold using annual data as shown in Figure 4. In 1997, the life cycle pattern of savings peaked at age group 55 to 59 years. The other deviation of the 1997 data is that the age groups 15 to 19 and 20 to 24 years saved more compared to the older age groups. Except for this, the annual data for 2000 and 2003 had younger age cohorts saving less and older age groups having higher saving rates.

Figure 4 **Household Saving Rate by Age Group of Head of Household (by FIES Years, 1997, 2000, 2003)**



Source: National Statistics Office, Philippines: Family Income and Expenditure Survey data, various years.

4. Econometric Models for Saving Rate

This section discusses the theoretical framework of the econometric model for saving rate. A set of possible explanatory variables are identified, based primarily on the life cycle model, and augmented with other determinants of saving rate as suggested by the available literature on saving.

The standard theoretical model that explains much of household saving behavior is the so-called intertemporal utility optimizing agent model (commonly known as the life cycle model). In this model, the household chooses its current consumption and savings, and an asset portfolio, so as to smooth its utility over time. There are several representations of this standard theory and the one provided here follows that of Coleman (1998). In this model, the household solves the following problem:

$$\text{maximize}_{C_t, A_t} \; E_t\left[U\left(C_t, C_{t+1}, \ldots, C_T; B\right) \middle| Z_t, Z_{t+1}, \ldots Z_T\right]$$

$$\text{subject to}: C_{t+s} + A_{t+s+1} = Y_{t+s} + (1 + r_{t+s})A_{t+s} \qquad \text{(i)}$$

where, C_{t+s} is the consumption at time (t + s);

Y_{t+s} is the income at time (t + s);

A_{t+s} is the asset held at the beginning of the period (t + s);

Z_{t+s} are variables affecting utility of consumption such as household size;

B is a bequest left to younger generation, $B \geq 0$;

r_{t+s} is the interest rate at time (t + s);

U is the utility function; and,

E_t is the expectation operator taken at time t.

The utility function U is assumed to be intertemporally additive,

$$U\left(\left(C_t, C_{t+1}, \ldots, C_T; B\right) \middle| Z_t, Z_{t+1}, \ldots Z_T\right) = \sum_{s=0}^{T} \beta^s v(C_{t+s}). \qquad \text{(ii)}$$

This particular assumption means that consumption is instantaneously enjoyed as it takes place and is independent of consumption at other times. The discount rate β indicates the patience of the consumer and is assumed to be less than one indicating that households prefer a marginal peso spent now than a marginal peso spent later.

The utility function $v(C_t,Z_t)$ is assumed to be quadratic in which households are risk-neutral. The model is known as the certainty equivalence model (Hall; 1978). Letting $v(C_t,Z_t) = v(C_t/\alpha(Z_t))$, where $\alpha(Z_t)$ is an equivalence scale for the demographic characteristics of the household, the solution to the maximization problem above equates marginal discounted marginal utility across time:

$$\frac{\partial v(C_t / \alpha(Z_t))}{\partial C_t} = E_t\left[\beta(1+r)\frac{\partial v(C_{t+1} / \alpha(Z_{t+1}))}{\partial C_{t+1}}\right] \qquad \text{(iii)}$$

The equation above implies that optimal consumption depends solely on the person's level of impatience and lifetime resource, the marginal propensity to consume from current income is the same as the marginal propensity to consume from expected future income (for instance, young people expecting higher income later in life will borrow against this income). In each period, the household plans to equate discounted consumption over time and in making that decision, what is important is the household's lifetime income and not current income. Moreover, when $\alpha(Z)$ varies over the life cycle, consumption varies accordingly--rising when there are children and falling as children become independent.

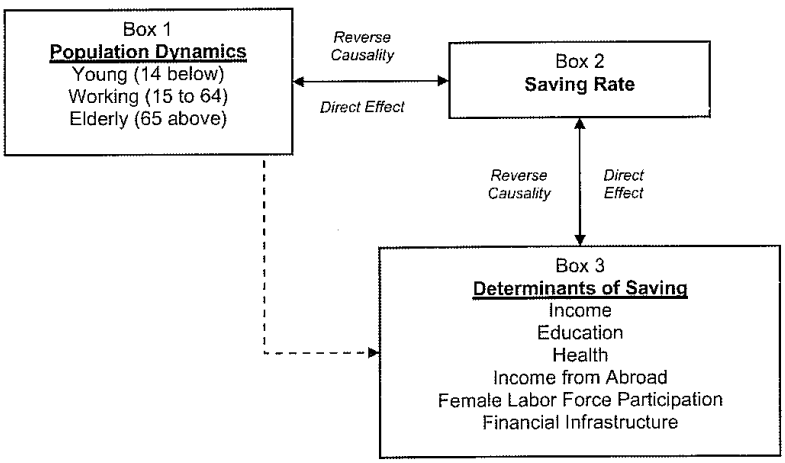
There are several shortcomings of the model above. For one, the elderly do not seem to dissave nearly as much as the model predicts. In fact, empirical evidence suggests that the elderly do save. Research suggests that rising longevity plays an important role in determining national savings. Lee, Mason, and Miller (1998, 2000) and Bloom, Canning, and Graham (2003) argue that the elderly do not dissave as much as what the life cycle model predicts, primarily because of the need to finance a longer period of retirement (the precautionary motive). These observations have been documented using household data from East Asia. It is therefore important that empirical models of aggregate savings include longevity as a determinant.

There is empirical evidence that liquidity constraints matter. For example, young people do not borrow against future income because they are concerned that they may not earn what they expect later in life (Coleman 1998). Also, there is an important issue of the bequest motive for saving and accumulating wealth. It is also important to understand why bequests are made since empirical evidence suggests that a large portion of accumulated capital stock results from bequests.

Despite the shortcomings of the life cycle model of saving behavior, it is still useful in providing a picture as to why people save. It seems that many individuals, especially the better educated, tend to smooth their consumption over their lifetime. Moreover, people do partly save for retirement. However, augmentation of the model to capture some of its shortcomings is needed and might be useful from the point of view of policy.

The framework of the econometric model is given in Figure 5. The econometric model estimates the direct effect of the population dynamics, particularly the impact of the young population (0 to 14 years) and the elderly population (65 years and above), on the household saving rate (the impact of the Box 1 on Box 2). At the same time, the model permits estimates of the effects of other determinants of saving rate (the impact of Box 3 on the Box 2). The reverse causality is represented by the arrow coming from growth (Box 2) going to the population dynamics (Box 1) and the other determinants of growth, notably education (Box 3). This reverse causality creates a problem in the estimation of the regression model, resulting in biased and inconsistent estimates. This problem is remedied through the introduction of instrumental variables into the regression equation.

Figure 5 **Theoretical Framework of the Econometric Model**



The basic econometric model for aggregate savings is the two-way error component fixed effects model where,

$$y_{it} = \alpha_i + \lambda_t + \underline{x}_{it}' \underline{\beta} + \varepsilon_{it} i = 1,2,...,14 \text{ and } t = 1,2,...,6$$

Here, y_{it} is the saving rate of Region i in time t, the vector $\underline{x}$ represents the determinants of saving discussed above, $\underline{\beta}$ is the vector of coefficients, α_i represents the regional and unobservable fixed effect, λ_t denotes the unobservable time effect and ε_{it} is the random error term assumed to be normally distributed with mean 0 and constant variance σ^2_ε. The coefficient $\underline{\beta}$ is estimated using the Generalized Least Squares (GLS).

A. Variable of Interest

The variable of interest in the econometric model of saving rate is the aggregate regional household saving rate from 1988 to 2003. The variable is defined as,

$$S_{it} = \left(\frac{I_{it} - C_{it}}{I_{it}} \right) * 100 \quad i = 1,2...,14 \qquad t = 1,2,...,6 \qquad \text{(iv)}$$

where S_{it} is the aggregate household saving rate of the i^{th} region at time t;

I_{it} is the aggregate (total) household income of the i^{th} region at time t; and,

C_{it} is the aggregate consumption (expenditure) of the i^{th} region at time t.

The data source for the aggregate regional household saving rate is the FIES of 1988, 1991, 1994, 1997, 2000, and 2003 (time periods). The data set consists of panel data with 6 time periods corresponding to the FIES years and 14 cross-sectional units equivalent to the 14 regions as defined in 1988. Although there are currently 17 regions, the geographical boundaries of the regions were kept constant throughout the period 1988 to 2003. The 14 regions as defined in 1988 are as follows:

Ilocos Region	Eastern Visayas
Cagayan Valley	Western Mindanao
Central Luzon	Northern Mindanao
Southern Tagalog	Southern Mindanao
Bicol Region	Central Mindanao
Western Visayas	National Capital Region (NCR)
Central Visayas	Cordillera Administrative Region (CAR)

B. Determinants of Aggregate Savings

In the construction of an econometric model for the saving rate, defined above, the following variables are identified as possible determinants.

1. Age Structure/Demographics

The average saving rate can be written as the sum of the saving rates of the different age groups in a population weighted by their income shares. That is, the average saving rate is,

$$\bar{s} = \sum_{i=0}^{T} s_i \frac{Y_i}{Y} = \sum_{i=0}^{T} s_i \frac{P_i}{P} \frac{y_i}{\bar{y}} \qquad \text{(v)}$$

This decomposition suggests that the age structure of the population matters. In theory (in the absence of a bequest motive), the dissaving of the elderly should offset the saving of the young so that in a stable population there will be no aggregate saving. However, as argued by Bloom, Canning, and Graham (2003), if the age structure of the population is unbalanced, which happens during a demographic transition, the saving behaviors of the various cohorts do not cancel out and aggregate saving (or dissaving) is expected.

The variables used to represent age structure (at the beginning of the period) are the youth share of the population and the elderly share of the population and are defined as follows:

i. Percentage of Young Dependents (aged 0 to 14 years), over the total population, at the beginning of the period; and,
ii. Percentage of the Elderly (aged 65 and above), over the total population, at the beginning of the period.

2. Level of Education

Empirical evidence suggests that schooling transition towards higher levels is one reason for the diverging saving patterns between Latin American countries (Mexico and Peru) and East Asian countries (Taiwan and Thailand). More educated individuals usually have higher incomes, and thus, higher saving capacity and this may be the potential factor behind differences in domestic savings.

iii. The percentage of household heads having at least high school diploma would be used to capture the effects of education on aggregate savings.

3. Female Labor Force Participation

Fertility and female labor participation are usually jointly determined, and they have a double effect on saving behavior--lower fertility rates imply fewer children in the average household, while higher participation implies more household members in the work force and thus, more income.

iv. To capture this effect, the percentage of women (15 years and above) in the labor force would be used.

4. Longevitiy

In addition, longevity raises the saving rates of every age group. This suggests that an addition of some function of life expectancy to the relationship is necessary.

v. Life Expectancy (at birth) in years (at the start of the panel period) would, therefore, be added in the econometric model as a possible determinant of aggregate savings.

5. Growth Tilting

Economic growth increases the relative income of the young and it not only increases average savings but also increases the effect of having a large young cohort. This leads to a phenomenon known as "growth tilting", making the impact of a large young cohort on savings larger in a fast-growing economy.

vi. To capture the effect of this determinant, annual average growth rate of per capita GDP (in 1985 prices) over the previous 5 years (in percent) of the period (example for 1988 panel, the average growth rate of GDP from1983 to 1987 was used) will be used.

6. Inflation

Periods of high inflation tend to be associated with highly negative real rate of interest and may deter opportunities for saving.

vii. Therefore, annual regional inflation rate (in percent) will be added in the econometric model as determinant of aggregate savings.

7. Presence of the Financial Infrastructure

Presence of financial infrastructure in the regions, such as number of banks, investment houses and other financial institutions can promote saving among households. However, problems encountered by these financial institutions such as closed banks can create a negative perception among the households and may be a disincentive for saving. In the econometric models, the researchers used appropriate proxies to measure the presence of financial infrastructure in the regions.

viii. The presence of financial infrastructure in the regions is measured using the average number of branches of banks in the region (using the average of three years: the FIES year, a year before, and after the FIES year); and

ix. The number of closed banks during the same three years was also included as a determinant of saving rate.

8. Initial Level of Income/ Initial Level of Income Growth

The magnitude of life cycle savings may depend on the region's income level (or initial income growth level) to capture the relationship of life cycle saving with the level of regional development.

x. The natural logarithm of the initial regional per capita GDP (measured in 1985 prices) shall be used to capture this relationship.

9. Remittance (Income transfer from abroad)

The model is also interested in looking at the contribution of remittances to the savings of the households. The data source for income transfer is the FIES.

xi. Percentage of income from abroad defined as aggregated household income (assistance) from abroad over total household income would be included in the model.

10. Institutional and Cultural Differences

The institutional and cultural differences across regions can be accounted for by allowing regional fixed effects in the estimation.

C. Empirical Analysis of the Model

The average regional household saving rate from 1988 to 2003 is 18.48% (Table 5). In 2003, the average household saving rate was even lower at

Table 5. **Summary Statistics for the Variables in the Econometric Model**

VARIABLE	Mean	Maximum	Minimum	Std Dev.
Saving rate	18.48	25.72	9.40	3.46
Log of initial income	9.15	10.33	8.49	0.44
Education	36.37	62.14	22.82	9.05
Percentage of young dependents	39.20	45.43	32.06	3.24
Percentage of Elderly	4.54	6.61	2.20	1.17
Log of Life Expectancy	4.18	4.26	4.09	0.04
Female Labor Force Participation	49.46	64.86	34.94	5.92
Household Income from Abroad (in %)	7.00	15.16	1.21	3.45
Inflation Rate	7.81	16.43	0.70	3.45
Number of Banks	410	2651	65	526
Number of Closed Banks	3.34	3.83	0.00	19.00

Source: Authors' computation using data from Asia Pacific Policy Center, Bangko Sentral ng Pilipinas, and National Statistics Office.

16.36%. The household saving rate of the Philippines pales in comparison with its neighboring countries in East Asia such as Taiwan and Thailand where the household saving rates were recorded at 30% for Thailand (in 1996) and 49% in Taiwan (in 1996).

To explain what drives household saving rate, an econometric model was built using panel data. The results of the two specifications using the Generalized Least Squares (GLS) are provided in Table 6.

Table 6 **Determinants of Regional Household Saving Rate (a)**
Dependent variable is aggregate regional household saving rate.
(Panel Data; Fixed Effects Model)

Variable	MODEL 1		MODEL 2	
	Coefficient	s.e. α	Coefficient	s.e.α
Log of initial income	5.40**	2.67	6.98*	2.89
Education	0.27*	0.15	0.27	0.16
Proportion of young dependents	-0.36**	0.17	-0.37*	0.21
Proportion of elderly	2.38***	0.56	2.12***	0.68
Female labor force participation	-0.04	0.09	-	-
Log of life expectancy	22.60	22.83	-	-
Percentage of Income from abroad	0.56***	0.23	0.44**	0.19
Inflation Rate	-	-	-0.20	0.20
Constant	-134.95	89.65	-54.76*	29.92
N	84		84	
Adjusted R-squared	0.69		0.7	

Notes: *** significant at 1%; ** significant at 5%; * significant at 10%;
α: standard errors are White's heteroskedasticity consistent
Source: Authors' computation using data from Asia Pacific Policy Center, Bangko Sentral ng Pilipinas, and National Statistics Office.

In Model 1 (base model), the determinants include initial per capita GDP, level of education, the demographic variables, female labor force participation, longevity variable (life expectancy), and the proportion of income from abroad. The two demographic variables have significant but opposite signs. On one hand, the percentage of young dependents has a negative and significant effect on saving rate which is consistent with the life cycle model and supports earlier studies that slowing population growth has been associated with high savings in East Asia (Harrigan 1998). On the other hand, the percentage of the elderly population has a significant but positive effect on saving rate. Under the life cycle model, in the absence of bequest motive, the elderly population should be dissaving. However, the data from the regional panel say otherwise. The result for the elderly population in the model runs in contrast with the result of the cross-country saving rate regression where it was found that the "presence of large proportion of elderly people in the population depresses saving rate, with the effect of the old being particularly large" (Bloom, Canning, and Graham 2003). The income and

education variables have both positive and significant effects on saving rate which are consistent with the earlier studies. The percentage of income from abroad is also positive and significant driver of saving rate, while the female labor force participation and measure of longevity are not significant.

Another variation of the model is given in Model 2 where inflation rate is incorporated into the regression model and the two insignificant variables (female labor force participation and life expectancy) excluded in the model. The results show that inflation rate does not play a significant role in determining aggregate saving rate.

A potential problem in the regression specification is the potential reverse causation from the saving rate to the level of income. Growth studies have shown that saving rate is a key variable in determining the speed of economic growth and the steady state level of income. There is also potentially a feedback from saving to education, implying that higher savings may give rise to higher levels of education. To solve the problem, we use instrumental variables, treating income and education as potentially endogenous. The instrumental variables used are initial geographical conditions (percentage of provinces in the region that are landlocked), percentage of households with access to electricity and measures of inequality.

Table 7 shows two specifications of the models using regression with the instrumental variables. In Model 3, the number of closed banks is incorporated to include presence of financial infrastructure (other variants of the model which also included the number of bank branches in the region was found to be insignificant).

Table 7 **Determinants of Regional Household Saving Rate (b)**
Dependent variable is aggregate regional household saving rate.
(Panel Data; Fixed Effects Model, Instrumental Variable)

Variable	MODEL 3		MODEL 4	
	Coefficient	s.e. α	Coefficient	s.e.α
Log of initial income	6.94**	2.71	6.98**	2.89
Education	0.27*	0.16	0.27*	0.15
Proportion of young dependents	-0.33**	0.16	-0.34**	0.14
Proportion of elderly	2.07***	0.61	2.03***	0.55
Percentage of Income from abroad	0.47***	0.17	0.51***	0.16
Inflation Rate	-0.21	0.20	-0.20	0.20
Number of Closed Banks	0.00	0.08	-	-
Constant	-53.93**	22.99	-53.99**	25.29
N	84		84	
Adjusted R-squared	0.7		0.7	

Notes: *** significant at 1%; ** significant at 5%; * significant at 10%;
α: standard errors are White's heteroskedasticity consistent

Source: Authors' computation using data from Asia Pacific Policy Center, Bangko Sentral ng Pilipinas, and National Statistics Office.

The results in Models 3 and 4 are somewhat the same as the previous two models with level of income, education, percentage of young dependents, percentage of the elderly and percentage of income from abroad as being significant drivers of saving rate.

The econometric model suggests that average per capita income is positively and statistically related to saving rate, supporting the life cycle hypothesis. The results from Model 4 (the preferred model) shows that a 100-peso increase in the average per capita income leads to an increase in the estimated mean saving rate by about 0.67 percentage point, all things being the same.

Education is an important determinant of saving rate, supporting earlier studies in other East Asian economies. A percentage point increase in the proportion of household heads with at least high school diploma increases the estimated mean saving rate by about 0.27 percentage point.

The results are mixed when it comes to the demographic variables, with the percentage of young dependents (ages 0 to 14 years) having a negative and significant impact on aggregate household saving rate while the proportion of the elderly (ages 65 years and above) has a positive and significant impact on the aggregate household saving rate. A percentage point reduction in the proportion of young dependents results in an increase in the average saving rate by 0.34 percentage point, while a percentage point increase in the proportion of the elderly results in an estimated increase of 2.03 percentage points in the average saving rate, all things being the same.

The study also confirms that remittance is one major source of aggregate household savings. A percentage point increase in income from abroad results in an increase in the estimated mean saving rate by about half percentage point, ceteris paribus. However, neither the rate of inflation (included in Models 3 and 4) nor the number of banks (included in Model 3) in the region plays a significant role in determining the aggregate saving rate.

5. Conclusion

The simulation results of the econometric model suggest that the Philippines' population dynamics play an important role in the aggregate household saving rate. A high proportion of young dependents (ages 0 to 14 years) creates a hindrance to the expansion of aggregate household saving, supporting the life cycle hypothesis on saving. The country, with its rapid population growth over the years, exhibits a big bulge in the lower portion of the age pyramid which resulted in a higher percentage of young dependents compared to the working population (ages 15 to 64). The swelling of the young dependency

group also increases the demand for education and health which results in lower levels of aggregate savings. This suggests that the country's unabated high population growth rate through the years have resulted in lower saving rates and consequently, lower economic growth rates.

Achieving a slower rate of population growth should be an explicit development objective of the government. Lower rates of childrearing will substantially increase the incentives for saving as experienced by East Asian countries like Singapore, South Korea, Taiwan, and Thailand.

An interesting result from the econometric model is the fact that the older population (ages 65 and above) still saves, contrary to the expectation of the life cycle model. This phenomenon has also been documented in studies in other Asian countries such as Taiwan and Thailand. Unfortunately, in the Philippines, as in other countries, there are cases when the elderly are being victimized by various forms of "investment scams". This suggests that efforts should be made to create awareness among the elderly regarding proper investment. The financial institutions should provide financial instruments that will fit the needs of the elderly population.

Education is also vital to increasing aggregate household savings. Policy-wise, efforts should be made to reduce the gender gap in education. These policies should encourage women to participate in the labor force.

Remittances are a major source of aggregate household savings as the econometric model shows. However, remittances can also be a disincentive for saving as these are withdrawn immediately to settle bills and to purchase consumer goods. Policy-wise, efforts should be made to encourage both the migrant workers and the remittance beneficiaries to save a portion of the remittances in secure financial intermediaries. Financial institutions should develop saving products that would be attuned to the needs of the lower income migrant workers.

Acknowledgments

The authors are grateful to Deputy Governor Nestor A. Espenilla, Jr. of the Bangko Sentral ng Pilipinas (BSP) for his support in this research on household saving. The authors benefited from discussions with Arsenio M. Balisacan and Ramon L. Clarete as well as the participants of the 11th International Convention of the East Asian Economic Association (EAEA) held in Manila, the participants of the roundtable discussions held at the School of Statistics, UP Diliman, and BSP for their comments and suggestions. The authors are thankful to Kristine Joy S. Briones for excellent research assistance and the Asia Pacific Policy Center (APPC) for providing the data used in the econometric models and an anonymous referee for comments and suggestions. Errors and omissions are sole responsibilities of the authors.

References

Attanasio, O.P. and Székely, M. (2001), "Household saving in developing countries: Inequality, Demographics and all that. How different are Latin America and South East Asia?" Working Paper No. 427, Inter-American Development Bank.

Bautista, R.M. and Lamberte, M.B. (1990), "Comparative Saving Behavior of Rural and Urban Households in the Philippines" Working Paper 90-15, Philippine Institute for Development Studies.

Bloom, D. and Canning, D. (2001), "Cumulative Causality, Economic Growth, and the Demographic Transition." In Birdsall, N., A. C. Kelly and S.W. Sinding, eds. *Population Matters, Demographic Change, Economic Growth and Poverty in the Developing World.* Oxford University Press.

Bloom, D.E., Canning, D., and Graham, B. (2003), "Longevity and Life-Cycle Savings." *Scandinavian Journal of Economics* 105: 319–38.

Bloom, D.E. and Williamson, J.G. (1997), "Demographic Transitions and Economic Miracles in Emerging Asia." Working Paper 6268, National Bureau of Economic Research.

Browning, M., Deaton, A., and Irish, M. (1985), "A Profitable Approach to Labor Supply and Commodity Demands over the Life Cycle," *Econometrica* 53-3: 503-543.

Browning, M. and Lusardi, A. (1996), "Household saving: macro theories and micro facts." *Journal of Economic Literature* 34: 1797-1855.

Coale, A.J. and Hoover, E.M. (1958), "Population Growth and Economic Development in Low-Income Countries." Princeton: Princeton University Press.

Coleman, A. (1998), "Household Savings: A Survey of Recent Microeconomic Theory and Evidence." Working Paper 98-8, New Zealand Treasury.

Deaton, A. S. (1992), "Understanding Consumption." Oxford: Oxford University Press.

Deaton, A. S. and Paxson, C.H. (2000), "Growth, Demographic Structure, and National Saving in Taiwan." In Chu, C. Y. Cyrus and Ronald Lee, eds. *Population and Economic Change in East Asia, Population and Development Review*, a supplement to Volume 26: 141-73.

Hall, R. (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy* 86(6): 971-987.

Horioka, C.Y. and Wan, J. (2006), "The Determinants of Household Saving in China: A Dynamic Panel Analysis of Provincial Data." Working Paper No. 12723, National Bureau of Economic Research.

Kelly, A.C. and Schmidt, R.M. (1996), "Saving, Dependency and Development." *Journal of Population Economics* Vol. 9: 365-86.

Kelly, A. C. and Schmidt, R.M. (2007), "Evolution of Recent Economic-Demograhic Modeling: A Synthesis." In Mason, A. and Mitoshi Yamaguchi, eds. *Population Change, Labor Markets and Sustainable Growth: Towards a New Economic Paradigm.* Elsevier.

Lee, R.D., Mason, A., and Miller, T. (1998), "Saving, Wealth and Population." Paper presented at the Symposium on Population Change and Economic Development at the Bellagio Center, Lake Como, Italy, November 1998.

_____. (2000), "Life Cycle Saving and the Demographic Transition: The Case in Taiwan." In Chu, C.Y. Cyrus and Ronald Lee, eds. *Population and Economic Change in East Asia, Population and Development Review*, a supplement to Volume 26: 194-223.

_____. (2001), "Saving, Wealth, and the Demographic Transition in East Asia." In Mason, Andrew, ed. *Population Change and Economic Development in East Asia: Challenges Met, Opportunities Seized.* Stanford: Stanford University Press.

Mapa, D. and Balisacan, A. (2004), "Quantifying the Impact of Population on Economic Growth and Poverty: The Philippines in an East Asian Context." In Sevilla, L.A., ed. *Population and Development in the Philippines: The Ties That Bind.* AIM Policy Center.

Mapa, D. S., Balisacan, A.M., and Briones, K.S. (2006), "Robust Determinants of Income Growth in the Philippines." *Philippine Journal of Development*, Volume XXXIII, Nos. 1-2: 1-32.

Mason, A. (2001), "Population and economic growth in East Asia." In Mason, Andrew, ed. *Population Change and Economic Development in East Asia: Challenges Met, Opportunities Seized.* Standford: Standford University Press.

_____. (2007), "Demographic Dividends: The Past, the Present and the Future." In Mason, Andrew and Mitoshi Yamaguchi, eds. *Population Change, Labor Markets and Sustainable Growth: Towards a New Economic Paradigm.* Elsevier.

Mason, A. and Lee, R. (2006), "Reform and support systems for the elderly in developing countries: Capturing the second demographic dividend." *Genus (Journal of Population Sciences)* Volume 62 No. 2: 11-35.

Modigliani, F. (1986). "Life cycle, individual thrift, and the wealth of nations." *American Economic Review* 76: 297-313.

Paxson, C. H. (1995), "Saving and Growth: Evidence from Micro Data." Working Paper No. 5301. National Bureau of Economic Research.

Radelet, S., Sachs, J., and Lee, J. (1997), "Economic Growth in Asia." *Emerging Asia: Changes and Challenges.* Asian Development Bank.

Rodriguez, J.A.A. and Meyer, R. L. (1988), "The Analysis of Saving Behavior: The Case of Rural Households in the Philippines." Working Paper 88-12, Agricultural Credit Policy Council.

Solow, R. (1956), "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70:65-94.

Swan, T. (1956), "Economic Growth and Capital Accumulation." *Economic Record* 32:334-361.

Williamson, J.G. and Higgins, M. (2001), "The Accumulation and Demographic Connection in East Asia." In Mason, Andrew, ed. *Population Change and Economic Development in East Asia: Challenges Met, Opportunities Seized.* Stanford: Stanford University Press.

World Bank (1999), "Why Do Savings Rates Vary Across Countries?" *World Bank Policy and Research Bulletin* Volume 10, No. 1:1-4.

Young, A. (1995), "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience." *Quarterly Journal of Economics* 110: 641-680.

Estimating Individual Health Expenditures Age Profiles from Household Level Data in the Philippine National Health Accounts[1]

Rachel H. Racelis[2], John Michael Ian S. Salas[2], Fe Vida N. Dy Liacco[3], and Racquel Dolores V. Sabeñano[3]

Abstract

The Philippine National Health Accounts (PNHA) is a framework for the compilation of information on the country's health expenditures and has been providing data important to health policy-making for more than a decade now. To maintain and further expand its usefulness, the PNHA underwent major restructuring in 2005 specifically in terms of health expenditure classification by uses of funds. The revised PNHA now includes a breakdown by age group of the beneficiaries of health care expenditures. The primary sources of data for household out-of-pocket (OOP) spending for health are the two surveys regularly conducted by the National Statistics Office (NSO), namely: the Family Income and Expenditure Survey (FIES) and the Annual Poverty Indicator Survey (APIS). These two surveys, however, only report health expenditure totals at the household level and actual health expenditures of individual household members are not known. Thus, household health spending attributable to specific groupings of household members, such as by age group, cannot be estimated directly from available survey data. To generate the age breakdown of household OOP health expenditures for the PNHA, a number of approaches were explored including: (1) the household per capita approach, (2) the simple regression approach, and (3) the use of related health information such as weights in a modified household per capita approach. Results from the application of these three approaches to data from a nationally representative household expenditure survey were generally found to be consistent – the age profiles all had the J-shape or U-shape, with higher mean spending for the very young and the elderly relative to the rest of the age groups. These overall shapes of the (indirectly estimated) per capita age profiles of household health spending were consistent with and validated by a reference per capita age profile that was computed from

1 A paper prepared for the Tenth National Convention on Statistics held on 1-2 October 2007 at the EDSA Shangri-La Hotel, Metro-Manila organized by the National Statistical Coordination Board. This paper was subsequently revised in April 2009 and included material that has became available after the conference.

2 Professor and Ph.D. Student, University of the Philippines, Diliman, Quezon City.

3 Statistical Coordination Officer VI and Statistical Coordination Officer IV, National Statistical Coordination Board, Makati City.

a small sample household health survey that collected health expenditure data at the individual person level. Furthermore, the indirectly estimated age profile is improved and was closer to the actual profile as more health related information is utilized in the estimation procedure (Mason 1987, 1988, 2001; Modigliani 1988; Lee and Edwards 2001; Mason and Lee 2004).

Keywords: National Transfer Accounts, economic life cycle.

1. Introduction

The PNHA underwent restructuring in 2005 to include a breakdown of household health expenditures by age of the health care beneficiaries. National household income and expenditure surveys in the Philippines however collect health expenditure data only at the household level. Various methods were explored to estimate the mean per capita age profiles of household OOP health expenditures using data available from these household surveys. The estimated per capita age profiles of OOP health expenditures were then used in the PNHA to estimate the age breakdown (of household health expenditures) – estimated mean per capita values were multiplied by the population size at each age to obtain total health expenditures attributable to each age. This paper describes three methods explored for estimating the per capita age profile of household OOP health expenditures in the PNHA and compares the results obtained from each of these methods.

What are per capita age profiles?

Per capita age profiles are intended to show average patterns of, for example, earnings and consumption expenditures of individuals over their life cycle. Individual age profiles basically constitute the set of mean per capita values at different ages, where the mean per capita value for an age group is computed (from sample survey data) by taking the sum of, say, income, expenditures or specific expenditure items such as health care that are attributable to the age group and dividing the sum by the number of sampled individuals belonging to the age group. The ideal data for estimating per capita age profiles would obviously be data collected for individual persons.

The larger context of the study of age profiles

Age profiles of consumption expenditures, income and other related economic variables (e.g., taxes paid) are studied for various purposes. Currently, the study of such age profiles are intrinsic to research on the economic consequences of population age structure change, more particularly population aging. Some of these researches include those that examine saving over the life cycle, demographic dividends and fiscal implications such as on taxable resources

and provision of public services (Mason 1987, 1988, 2001; Modigliani 1988; Lee and Edwards 2001; Mason and Lee 2004).

The age profiles for income and consumption are used to determine the age groups that are economically dependent in a given population (i.e., age groups consuming more than they are earning) and to estimate the amount of resources required for their support. Furthermore, such analyses have been used to examine up to what extent dependent population needs are supported by family resources and what additional support systems (such as public provision) are needed. One set of studies that have recently emerged in this area of research used the National Transfer Accounts (NTA) which provides a comprehensive approach to estimating age profiles of consumption expenditure components, income, and asset accumulation among others, and to measuring all inter-age transfers at the aggregate level (Maliki 2004; Mason, Lee, Tung, Lai, and Miller 2005; Clark, Ogawa, and Matsukura 2007; Ogawa, Mason, Maliki, Matsukura, and Nemoto 2007; Racelis and Salas 2007; Ladusingh and Narayana 2008).

The methods applied in the NTA, including those for estimating per capita age profiles of household health expenditures, are described in a draft NTA manual (Mason, et al. 2009). The various methods described in the NTA manual had been developed in the course of applying NTA in many countries around the world – different methods to suit varying situations of data availability. The NTA has been applied in research done in 27 countries in Asia/Pacific, the Americas, Europe, and Africa.

The interest to study income and consumption patterns by age is expected to become more important in the near future as more countries experience rapid changes in their population age structure, more specifically aging of population. More national governments will be needing information to be able to formulate as well as implement appropriate and timely measures to contend with the economic impacts of population change. In the case of health care, for example, there is a need to determine the scale and timing of aggregate health expenditure increases as the proportion of the elderly in the population rises.

The PNHA and household health expenditures age profile

National Health Accounts (NHA) constitute a systematic compilation of health expenditure data for a country. In the Philippines the PNHA is produced annually by the National Statistical Coordination Board (NSCB). The PNHA was established in 1997 and has since then been useful for health policy making and monitoring. The design of the PNHA was revised in 2005 to expand its usefulness.

The revised PNHA now includes a table showing national health expenditures with a breakdown by age of beneficiary of health spending (Racelis, Dy Liacco, Sabeñano, Beltran, and Manaog 2006). This expenditure breakdown by age is shown in the PNHA for every type of payor for health services. Payors for health services in the Philippines include the national and local governments, Philippine Health Insurance Corporation (PhilHealth), private health insurance providers, employer-provided schemes, households (out-of-pocket or OOP for health care), and others such as non-profit institutions. This paper focuses only on showing how the age profile of household health expenditures may be estimated.

The primary sources of household OOP health expenditure data in the Philippines are the two household surveys regularly conducted by the NSO. FIES is conducted every three years, for example in the years 1997 and 2000, and the APIS is conducted during non-FIES years, starting in 1998. These two surveys collect health expenditure data only at the household level. That is, health expenditure data on individual household members are not available. Thus, the entries for household OOP health expenditures (column) in the new PNHA table showing the age breakdown cannot be computed directly from data in either the FIES or the APIS.

This data constraint in the Philippines is not uncommon. In the application of NTA in many countries, it was recognized that most expenditure data, including health expenditure data, are collected in surveys at the household level rather than at the individual level (Lee, Lee, and Mason 2004). Thus, various methods for estimating individual age profiles from household level expenditure data have been explored and tested such as those suggested in the NTA manual.

The minimum data needed from a household survey to be able to assign household level health expenditures to different age groups are (1) complete rosters of members for all households, and (2) age data for every household member. The method of allocating household health expenditures to age groups can be improved if there are additional health-related information on household members available from the survey, such as health/illness status and health care utilization. The additional health-related data can indicate each household member's consumption of health services and household health expenditures can then be assigned to each of the household members more appropriately.

The objective of this paper is to demonstrate feasible methods for estimating individual per capita age profiles from household level data for household OOP health expenditures. It focuses on three methods suggested in the NTA manual that were explored to estimate the age profile of household health

expenditures in the revised PNHA. The choice of methods tested was driven mainly by limitations in the data.

But how do the age profiles estimated using the different methods in fact compare with the true age profile? Data from the 1991 Philippine Institute of Development Studies-Department of Health (PIDS-DOH) Household Survey is used to establish the reference age profile for this assessment (i.e., a reference age profile as an approximation of the true age profile). The survey data was collected for the year 1991 and is unique because it contains health expenditure data at the individual person level. The reference age profile is estimated from individual level data.

Sections of the paper

The three methods explored for estimating the per capita age profile of household health spending for PNHA are discussed in Section 2. The resulting age profiles from applying the three methods to the 1999 APIS data, a nationwide household survey, are presented in Section 3. Section 4 presents the reference per capita age profile computed from the small sample survey containing individual level health expenditure data. Section 5 provides a summary of the findings and the conclusion.

2. Approaches to Estimating Individual Level Expenditures from Household Level Data

The first method uses only the minimum data possible – i.e., data for every household on household health spending and ages of its member. The second method may be applied using only the minimum data or with some additional health-related data if these are available. The third method involves the use of additional health-related data that may be available from the survey or from secondary sources.

Household per capita approach

In this approach (Racelis, et al. 2006; Mason, et al. 2009), the age profile of household health spending is generated in two steps. First, the total health expenditure of a household is divided by the total number of household members. The computed value for household per capita health spending is then assigned as an estimate of the individual spending of every member of the household. The simplifying assumption being made here is that all household members had incurred (equal) health expenditures during the survey year.

Second, when the individual health expenditure has been computed for all households (i.e., Step 1 completed for all households), all individual persons or members from all households are then pooled together and sorted by age. Then the average health expenditure for every single year age group using expenditures assigned in Step 1 to individuals is computed.

Simple regression approach

In the simple regression approach (Mason, et al. 2009), a purely demographic linear regression equation is estimated using ordinary least squares (OLS). Household health expenditure H is the dependent variable, the number of household members n_j at each age j (in single years) are the independent variables and the intercept is set to zero as follows:

$$H_i = \beta_0 n_0 + \beta_1 n_1 + \beta_2 n_2 + \ldots + \beta_p n_p + e_i, \qquad (1)$$

such that $e_i \sim \mathrm{N}(0,\sigma^2)$. The estimated regression coefficient β_j is the estimated average health expenditure per person for individuals age j years. That is, the estimated regression coefficients together constitute the age profile for OOP health spending.

The above regression specification uses the minimum data possible. The NTA manual however suggests this method only as a last resort. The manual also suggests using the regression method with additional data on health service utilization of household members. Instead of number of household members by age as explanatory variables, the number of household members using health services by age are used. The specification can be further modified to differentiate situations of households by including household income, health insurance coverage and other variables as explanatory variables (Racelis, Russo, and Mason 2003). The modified regression approach is feasible only if the data on additional variables are available in the survey dataset.

Modified per capita approach

In this approach the household per capita approach is modified by assuming varying shares instead of equal shares of the household health expenditures assigned to different household members (Racelis and Salas 2007). The age profile of household health spending is generated in three steps. First, the relative shares or weights for allocating household health expenditures to household members (of different ages) need to be obtained. The per capita age profile of health spending of another country with a similar health profile may be used as the age-specific weights. Or the relative weights can be generated using country data on health-related, non-expenditure variables. The age profile shown in this paper was estimated by taking the latter

approach in which health facility utilization rates by age were used as weights (Racelis and Salas 2007).

Second, as in the household per capita approach, the household health spending (H) is allocated to its members, but this time using the utilization rates f^k as weights to compute for the health expenditures H^k of household member with age k. The health expenditures of an individual household member of age j is estimated as follows:

$$H^j = H \times \frac{f^j}{\sum_j f^j} \qquad (2)$$

Third, as in the household per capita approach, after the assignment of household health expenditures to individual household members has been completed for all households, health spending of each member of all households are then pooled together and sorted by age. Average health expenditures at every age can then be computed.

3. Applications Using the 1999 Annual Poverty Indicator Survey (APIS)

This paper presents results from previous research for the PNHA that used the 1999 APIS, which was then the most recent household expenditure survey dataset available. For purposes of the PNHA, the household per capita, the simple regression and the modified household per capita methods were explored and the resulting age profiles generated are shown in Figure 1. The modified regression approach was not used because of data constraints.

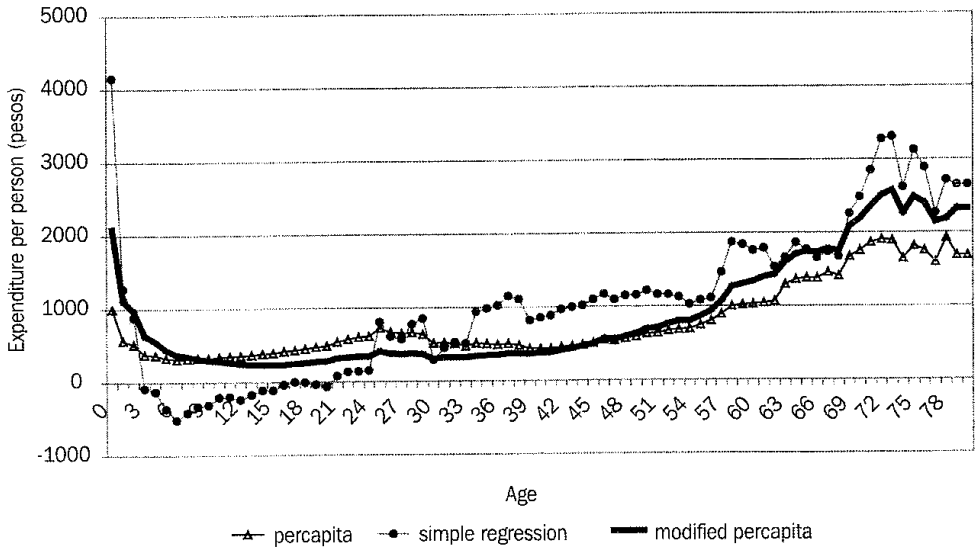
In general, the estimated age profiles of household health expenditures generated using all three approaches have J-shape or U-shape curves, where the per capita health expenditure for very young children and the elderly are generally higher than the rest of the population.

The J-shape or U-shape profile is captured in the results of the household per capita and the simple regression approaches because households with members who are "expensive" health-wise (i.e., the very young and the elderly) generally tend to have higher OOP spending for health. In the household per capita approach, the computed per capita health spending is higher for such households on average; and, on the contrary, (on average) lower for households where there are neither very young nor elderly members. In the simple regression approach, the estimated regression coefficients showed that the increase in the health expenditure of a household is higher with the addition of very young or elderly individuals as members of the household.

The age profile from the household per capita approach tends to be flatter compared to the other two profiles because of the "averaging" effect. The simple regression approach over-exaggerates the "smallness" of health spending for older children and teen-agers, presenting negative coefficients.

The household per capita and the simple regression approaches use only two sets of information to estimate the per capita health expenditure profile by age (i.e., total household spending for health and age composition of household members). The modified per capita approach, in addition to using total household health spending and members' age profile, incorporates information on differences in intensity of use of health services among persons of different ages. The per capita age profile for the third method is observed to lie between those based on the household per capita method and the simple regression method.

Figure 1 **Health Expenditure Per Person By Age: 1999 Annual Poverty Indicator Survey**



Source: Authors' computations using the 1999 Annual Poverty Indicator Survey.

4. The Reference Household Health Expenditures Per Capita Age Profile

The J-shape or U-shape patterns observed in the three estimated per capita age profiles for household expenditures in Section 3 may be assessed for validity by comparing these with the the age profile pattern generated directly from actual individual level health expenditures data. But as stated previously, household expenditure data have rarely been collected at the individual level in the Philippines and in other countries; thus, such data even if not recent, are most valuable. For purposes of this paper, the 1991 PIDS-DOH Household Survey was used to establish the reference household health expenditures per capita age profile. The survey covered 2,798 households in four provinces (with 14,227 individual household members) and collected household health expenditure data at the individual household member level.

The per capita age profile for household health expenditures was computed directly from the survey data by simply taking the total health expenditures of the sampled individuals by age and dividing the totals by the number of individuals in each age group. This actual per capita age profile is shown in Figure 2. (Note that the number of age groups has been reduced because of the relatively smaller sample size compared to the APIS.) In addition, the age profiles that would result from the household per capita method and the modified per capita method were simulated using data from the same survey by pooling individual health expenditures at the household level, and then applying the two methods as described in Section 2. The simulated age profiles could be compared directly with the actual profile.

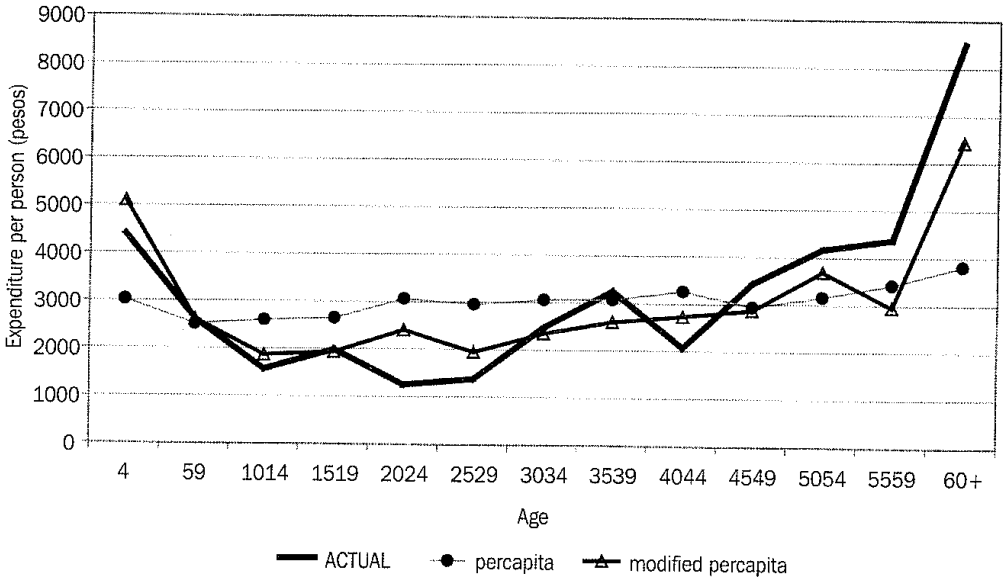
Examining the age profiles in Figure 1 and Figure 2, observations include the following: (1) the J-shape or U-shape pattern of per capita health expenditures over the life cycle in Figure 1 is validated by the actual per capita profile in Figure 2; (2) the household per capita approach shows the flattest profile in both Figure 1 and 2; and (3) the modified per capita method improves on the age profile derived from the household per capita method by pushing the profile closer to the actual (Figure 2).

5. Summary and Conclusion

The estimated Philippine age profiles for household OOP health expenditures generated by applying three different approaches to data from a nationally representative household survey have the J-shape or U-shape pattern. These overall patterns are validated by a per capita age profile computed from actual individual level data for household expenditures. The profile based on the household per capita approach is generally flatter compared to those

generated using the other two approaches for older children and teen-agers, and is recommended as a last resort. The age profile generated using the modified per capita approach (with health facility utilization rates as weights for household allocation of health expenditures) is, in general, situated between the profiles from the household per capita and simple regression methods.

Figure 2 **Health Expenditure Per Person By Age: 1991 PIDS-DOH Household Survey**



Source: Authors' computations using the 1999 PIDS-DOH Household Survey.

Given alternative approaches for indirectly estimating the age profile of OOP health spending, how should one choose which one to use? One important rule to follow in the selection of the approach is to use that method that would fully utilize every piece of data available. With the use of more information, the estimated age profile would be closer to and be a more realistic approximation of the true profile, as demonstrated by the age profiles presented in Section 4.

5. References

Clark, R., Ogawa N., and Matsukura R. (2007), "Population Aging, Changing Retirement Policies and Lifetime Earnings Profiles in Japan," in *Population Aging, Intergenerational Transfers and the Macroeconomy*, Robert Clark, Naohiro Ogawa, and Andrew Mason, Cheltenham, UK, and Northampton, MA: Edward Elgar.

Ladusingh, L. and Narayana, M.R. (2008), "Population Aging and Life Cyle Deficit: The Case of India," In *Population, Poverty and Health-Analytical Approaches*, K.K.Singh, R.C.Yadava and Arvind Pandey. New Delhi: Hindustan Publishing Corporation.

Lee, R. and Edwards, R. (2001), "The Fiscal Impact of Population Change." In *Seismic Shifts: The Economic Impact of Demographic Change.* J.S. Little and R.K. Triest, Federal Reserve Bank of Boston Conference Series, 46.

Lee, R., Lee S. H., and Mason, A. (2004), "Individual Earnings and Consumption Profiles: What Do We Know?" Symposium on Population Aging And Economic Productivity, Vienna Institute of Demography, December 2004 (November 29, 2004 draft, www.ntaccounts.org).

Maliki (2004), Estimation of Intergenerational Education Transfers in Indonesia, December 2004. www.ntaccounts.org.

Mason, A. (1987), National Saving Rates and Population Growth: A new Model and Evidence. In *Population Growth and Economic Development: Issues and Evidence.* D.G. Johnson and R.D. Lee Social Demography Series. Madison, Wis. University of Wisconsin Press.

_____. (1988), "Saving, Economic Growth and Demographic Change." *Population and Development Review*, 14(1).

_____. (2001), Population Change and Economic Development in East Asia: Challenges Met and Opportunities Seized. Stanford: Stanford University Press.

Mason, A. and Lee, R. (2004), "Reform and Support Systems for the Elderly in Developing Countries: Capturing the Second Demographic Dividend". International Seminar on the Demographic Window and Healthy Aging: Socioeconomic Challenges and Opportunities, China Center for Economic Research, Peking University, Beijing.

Mason, A., Lee, R., Donehower, G., Lee, S.H., Miller, T., Tung, A.C., and Chawla, A. (2009), *National Transfer Accounts Manual (Draft Version 1).* National Transfer Accounts Project, East West Center, Honolulu and University of California (Demography and Economics Center), Berkeley, www.ntaccounts.org.

Mason, A., Lee, R., Tung, A.C., Lai, M.S., and Miller, T. (2005), "Population Ageing and Intergenerational Transfers: Introducing Age into the National Accounts." Forthcoming. In *Developments in the Economics of*

Aging of Aging, David Wise, National Bureau of Economic Research and University of Chicago Press.

Modigliani, F. (1988), "Measuring the Contribution of Intergenerational Transfers to Total Wealth: Conceptual Issues and Empirical Findings." In *Modelling the Accumulation and Distribution of Wealth.* D. Kessler and A. Masson, Oxford: Oxford University Press.

Ogawa, N., Mason, A., Maliki, Matsukura, R., and Nemoto, K. (2007), "Population Aging and Health Care Spending in Japan: Public and Private Sector Responses." In *Population Aging, Intergenerational Transfers and the Macroeconomy*, Robert Clark, Andrew Mason, and Naohiro Ogawa, Cheltenham, UK, and Northampton, MA:Elgar Press.

Racelis, R., Dy Liacco, F., Sabeñano, R., Beltran, M., and Manaog, T. (2006), "The National Health Accounts of the Philippines: Continuing Development and New Findings." *Philippine Journal of Development*, Vol. 33, No. 61.

Racelis, R., Russo, G., and Mason, A. (2003), "Baseline Estimates of Health Care Expenditures of Older Persons: Philippines, 1994." Honolulu: East West Center, Aging Populations: Health Systems and Policy Reform Project.

Racelis, R., and Salas, J.I. (2007), "Measuring Economic Lifecycle and Flows Across Population Age Groups: Data and Methods in the Application of the National Transfer Accounts (NTA) in the Philippines." Makati City: Philippine Institute for Development Studies, Discussion Paper Series No. 2007-12.

Estimating Food Poverty Directly from Consumption Data, Comparison with Official Method, and Suggestions for Change[1]

Isidoro P. David, Wilma L. Molano, Emma A. Fabian, Zita V. Albacea, and Rechel G. Arcilla [2]

Abstract

A quality assessment of the basic data from the two main sources of food consumption, Food Consumption Survey (FCS) and Family Income and Expenditures Survey (FIES), is made. National Statistical Coordination Board's (NSCB) official method of estimating food poverty incidence is not a direct user of food consumption from these sources, but relies instead on prescribed menus to compute food poverty lines. The method performed poorly when validated using FCS data. Alternative methods are discussed and tried using FCS and FIES data. Two are money-based methods that compute a food poverty line and compare this with either per capita income or expenditure. Another more promising group is composed of non-money based distribution-free methods, relying mainly on the food consumption converted into energy and other nutrients. One fits an empirical cumulative distribution function of per capita dietary energy consumption and uses this to estimate the food poverty incidence for any choice of energy threshold. A second eschews the difficulties and distortions associated with per capita calculations and rely instead on a comparison between the family's total consumption and its recommended energy and nutrient intake. The numerical results point to much higher food poverty incidences than generally perceived, the latter conditioned perhaps by official estimates. The need for empirical research to find better data capture methods for collecting food consumption data and for validating the numerical results from this study is highlighted.

Keywords: Food poverty, dietary energy consumption, cost of basic needs.

1 Presented at the Philippine Statistical Association (PSA) Anniversary Conference, 19 September 2007, PSSC Building, Diliman, Quezon City. Some results here were taken from a power point presentation entitled Food Poverty Estimation and Monitoring Surveys: Analyses and Suggestions for Improvement, by the same authors, at the Philippine Economic Society 44th Annual Meeting on 21 November 2006, Bangko Sentral ng Pilipinas.

2 President, PSA; Supervising Science Research Specialist, FNRI; Chief, Income and Employment Statistics Division, NSO; Associate Professor, University of the Philippines at Los Baños; and Assistant Professor, De La Salle University. We are especially grateful to the Food and Nutrition Research Institute and National Statistics Office for putting at our disposal the files from their 2003 Food Consumption Survey, Family Income and Expenditure Survey, and Labor Force Survey.

1. Introduction

Being poor have many dimensions and nuances. Lack of money, hence of the goods and services that money can buy, comes to mind first. Global achievement of the World Bank's vision, "A World Free of Poverty", which is embraced by most international development agencies, is monitored mainly through the number of people who live on less than one dollar a day (extremely poor) and sometimes on two dollars a day (poor). One can also be poor in opportunities directly or not directly associated with making more money, e.g. in choosing where to live, study, get health care, work and chances of finding gainful work, what we can say or write, and practice of our faiths. Being poor can be linked to a sense of insecurity about physical safety and about food availability in the next six months. In the extreme case, we can go otherworldly and talk about being poor in spirit. However, the need to determine who is poor, how poor, and where (s)he is brings us back to what can be observed or measured *in situ.* Difficulties in proposing a comprehensive definition of the state of being poor that is acceptable to the majority of stakeholders in society compels us to consider still narrower definitions.

Assessing poverty through *poverty lines* is based on a narrow definition of being poor as not earning (or spending) enough to provide for (a) the minimum amount of nutrition needed to sustain normal physical activity and good health and (b) essential non-food requirements such as clothing, shelter, primary schooling, basic health care, and the like. Most poverty line estimation methods involve estimating (a) first, then adjusting the estimate for expenses to meet (b). That is, total poverty line (tpl) = food poverty line (fpl) + non-food poverty line (nfpl) where the acronyms stand for total poverty line, food poverty line and non-food poverty line, respectively. Those who earn less than tpl are called *poor.* Those who earn less than fpl are called *food poor,* which the NSCB also calls subsistence poor or core poor. There are two features of the poverty line methods that we would like to mention because of their relevance to the main ideas of this paper. One is that everything is valued in money terms. There are, however, alternatives not dependent on any money metric. Another is that tpl and in some cases nfpl, are derived as functions of fpl; therefore, getting the fpl estimation method right is crucial to the entire poverty assessment exercise.

The choice of data capture influences the underlying conceptual and operational definitions, hence the interpretation and comparability also of the poverty estimates. The methods covered in the other paper in this conference (Maligalig 2007) use what may be described as self-assessed poverty, where the respondents are asked questions such as, whether they have experienced not eating or not having enough to eat during the prescribed reference period because there was no food available or there was no money to buy food; or

whether or not the respondent feels that the household is unable to eat what is otherwise a more balanced diet. That paper describes the estimates from these surveys as *direct* measures because they were computed directly from the respondents' self assessment of their poverty situation. This method of data capture offers a number of advantages. The questions can be phrased to address broader food inadequacy concepts, such as insecurity about food availability during the reference period, or of very narrow concepts such as hunger, which may be associated with persistent lack of food during the reference period. Intra-family analysis is possible by addressing the questions to different members of the family. There are disadvantages too, however. Lack of comparability of the estimates can result from potential culture- or group-confounded responses to the same questions.

FCS, food budget surveys and FIES generally use other methods of data capture. In the Philippines, the National Statistics Office's (NSO) FIES and the Food and Nutrition Research Institute's (FNRI) FCS are the two regular sources of family food consumption data disaggregated into quantities and prices by item (e.g., rice, pork, mungbean, etc.) Empirical comparisons of the basic data from these sources are presented in the next section.

The rest of the paper deals with the food poverty part of total poverty only; i.e., fpl in tpl = fpl + nfpl in the case of money-based measures. Section 3 analyzes the official NSCB method of computing fpl and the concomitant food poverty incidence estimates. The surprise is that the official method does not use either FIES or FCS family level food consumption data directly. The ability of the method to correctly classify families into food-poor and not food-poor classes is tested empirically using the FIES and FCS samples. Sections 4 and 5 try to answer a number of *what ifs*: What if the better of the FIES or FCS data are used to estimate food poverty statistics? Two methods involve computing fpls and comparing these with per capita income estimates also from FIES. A third method does away with money measures altogether (prices and income), and assesses food poverty through the empirical cumulative distribution function of nutrient consumption. A fourth avoids problems that accompany per capita calculations. Section 6 summarizes the main findings and recommendations.

2. Sources of Food Consumption Data

2.1 Household Surveys Integrated through a Master Sample Design

What some might consider gold standard for measuring human food consumption would involve recording the individual's food intake over a period of time, say one day or week or month. This is not done for many

reasons, including that family members eat together. In fact, the conventional survey definition of a household is a person or group of persons who live and prepare and consume food together. And a family consists of household members related by blood or marriage, i.e., household minus helpers and boarders. Thus, except for traits like height and weight that can be measured easily on each person, food consumption and even income and expenditure are recorded as family totals. This is the case with the two main sources of food consumption data in the Philippines, namely, NSO's FIES and FNRI FCS. FIES is done every three years, the last one in 2006. As of this writing, the 2003 round is the last whose results and public use files had been released. FCS is a module in FNRI's National Nutrition Survey (NNS) that also has anthropometric, clinical and biochemical surveys. NNS is conducted every five years, the last one being in 2003 also.

From 2003, NSO has used a master sample for all its household or family-based surveys, including FIES and the quarterly Labor Force Survey (LFS). The master sample consists of some 50,000 sample families divided into four equal-sized independent replicates, also called interpenetrating sub-samples by their original proponent, Mahalanobis (1946). To ease respondent burden, the master sample design features include partial replacement of sampling units from one year to the next; thus, there is a complete match of samples within the year and partial match between years (NSO, Master Sample Documentation). The 2003 NNS sample is one randomly chosen replicate from the NSO master sample. (FNRI 2005). This sharing of samples, coupled with the use of consistent and unique sampling unit codes down to the family level, provide users the opportunity for more comprehensive analyses by linking and merging databases of different surveys. This possibility is exploited in this study. We used matched samples from the 2003 FCS, FIES, and LFS.

2.2 About FIES

The 2003 FIES had 42,094 responding families, each visited twice. The first visit in July had the first six months of the year as reference period and the second in January 2004 covered the remaining six months. Data capture was by face-to-face interview of any responsible adult member of the family present at the time of visit who, in the judgment of the interviewer, is knowledgeable about the subject covered in the questionnaire. The latter is 70 pages long, 16 of which are on food consumption. Although the reference period for each visit is six months, what is actually asked is average weekly consumption and expenditure by food item (e.g., rice, pork, etc.).[3] The quantity consumed, unit

3 The question was phrased as: "During the period specified, did you ... consume ...(food group, e.g. fruits and vegetables?" If yes, "On average, how much is your weekly consumption of the following?" (list). Note that period specified is previous six months prior to interview.

price and value consumed for each item are in the questionnaire. However, since the thrust of FIES is estimation of income and expenditure, and since the NSCB official methodology for estimating food poverty incidence uses these two estimates only (see section 3 below), we think it is fair to say that as far as food consumption is concerned the emphasis in FIES is more on the family's total food expenditure than on quantities and unit prices of the food items consumed. Thus, much less attention is given to quantities and unit prices of the individual commodities, to the point that in some FIES rounds these had not been encoded, much less analyzed.[4] This can only impact negatively on the quality of the data on food quantities and unit prices, as was borne by a comparison with corresponding data from the FCS reported in sub-section 2.4 below.

Every food consumption survey is confronted with accounting for family members who eat out during the reference period. With FIES, the value of the food eaten by such members is estimated and added to the total food expenditure, as well as food purchased for special occasions. However, no adjustment is done in the case of guests who break bread with the family.

One advantage of surveys sharing the same sample is that variables of common interest, e.g., demographic characteristics, may be covered in one questionnaire only. In this study using food consumption data from FIES and FCS, the age-sex composition of the sample families were extracted from the 2003 LFS.

2.3 About FCS

The 2003 FCS employed a radically different data capture method (FNRI 2005). In place of enumerators, professional nutritionists visited the sample families and weighed separately all the food ingredients that were cooked and consumed from dawn till everyone had gone to bed in the evening; weighed the portions saved, fed to pets, thrown away; etc; same-day recall interview was done to estimate the amount of food from meals taken outside the home; and guests' consumption was estimated and netted out. The prices of the food items were asked from the family on the same day and the costs of food not bought (e.g., own produced or given) were imputed. The meal pattern of each family was recorded also; i.e., the usual number of meals and snacks taken in a day.

4 Encodings of the food consumption part of FIES were made for special studies: 2000 FIES for ADB- sponsored studies under the same ADB project that helped develop the 2003 master sample, and for FAO for a special study on malnutrition; 2003 FIES for an ongoing NSO-FAO project on food insecurity and poverty, and of the matched FIES-FCS sample for this paper.

It is easy to be convinced that food weighing yields very accurate results by drastically reducing measurement errors. However, very few countries still use the method because it is very costly, technically demanding, and it requires getting prior permission of the sample families.[5] For these reasons, the FCS module of the 2003 NNS was reduced to 3,300 sample families representing a random subsample of one replicate of the NSO master sample. Also, the food weighing was for one day only, with the sample families distributed equally and at random among the seven days of the week. With limited number of FNRI permanent professional nutritionists and field survey specialists, the data collection which started in July was completed in December 2003.

2.4. Comparison of FIES and FCS Data

2.4.1 At Item Level

There are close to 150 items each in the food lists of the two sources. The annual per capita consumption of these items can be estimated and compared. The results for selected items and from the common FIES – FCS sample are shown below (Table 1).

For staples like rice and corn, the FIES and FCS estimates are reassuringly close.[6] The other items in Table 1 seem to indicate that FCS yields higher values than FIES.

In relative terms, the estimates begin to diverge significantly as items become less frequently eaten, which upon reflection is to be anticipated. On one hand, staples are taken almost daily in portions within a narrow range (i.e., there is a limit to how much rice an individual can eat in one day) and most families would be reporting non-zero values even with one-day weighing only in FCS. The "usual one week consumption" reference in FIES will also be expected to lead to most reliable results for staples that the family spends on and eats regularly and most frequently. Thus estimates for these kinds of variables can be expected to be accurate and with small sampling

5 Some of the other methods used are face-to-face interview with recall periods ranging from daily to up to a year, diary method where the family keeps a record of food items bought and consumed for a period ranging from a month to a year, and diary method complemented by weekly visit by survey personnel to check on the diary entries. The NSO method is a variant of the first where what is asked is "usual weekly consumption".

6 Until 2000, the Bureau of Agricultural Statistics (BAS) conducted ad-hoc food consumption surveys mainly to estimate per capita consumption of staples. Annual per capita rice consumption estimates from the 1995 and 2000 surveys were 103kgs and 105kgs respectively. (From 16 August 2006 email communication with BAS Director Romeo Recide.)

Table 1. Annual Per Capita Consumption of Selected Food Items in Kilograms

ITEM	FIES	FCS	(FIES-FCS)/FCS (per cent)
Rice	104.0	107.0	-3
Corn	10.6	10.6	0
Fresh Chicken	4.1	6.6	-38
Fresh Beef	1.2	1.8	-35
Fresh Pork	6.0	11.6	-48
Banana	4.6	9.1	-49

Source: Family Income and Expenditure Survey, Food Consumption Survey (2003).

errors. On the other hand, non-staples that are eaten infrequently may be considered rare events; i.e., on a given day these will have zero values for most families and non-zero for a few only. Such variables will have much higher sampling errors. It is also not difficult to imagine that respondents will have a harder time "guessing" relatively accurately the amount and cost of their usual one week consumption of such food items.

In other words, FCS should provide more accurate amounts than FIES for foods that happen to be eaten during the one day weighing for practically all items, but more particularly for those that are not bought by standard weights (e.g., native cakes, eggs, watercress or 'kangkong' in which only the top portion is cooked); when what is bought is generally not consumed within the day or week (e.g., cooking oil, soy sauce, sugar); and when not all that go into the pot is consumed. On the other hand, FCS registers a zero value (hence a zero estimate) for any item not consumed by the sample family during the one day food weighing, but which is most likely consumed in the course of the year (hence the true value is not zero).

Thus, after the main staples, further item-wise comparison of the two data sets is unlikely to lead to definitive conclusions. The following subsection deals with comparisons based on more stable non-zero estimates of nutrient contents of the food consumed in one day.

2.4.2 At Nutrient Level

The FCS recorded that *44%* and *21%* of its sample had members that ate out and had guests who dined with their host families an average of 2 meals respectively, during the food weighing day. These surprisingly large proportions suggest that failure to account for their shares to the total family consumption could affect estimates significantly. While both surveys make adjustments for outside meals, only FCS estimates net out the portions taken by guests.

FNRI has developed two key tables required for nutritional and dietary assessment. One is a computerized food composition table that converts food items expressed in original standard units (e.g., kilograms) into corresponding nutrient contents (e.g., kilocalories of energy, grams of protein). The consumption records of the matched sample families from FIES and FCS were put through this conversion process and used in this study. The FCS x FIES scatter plot of the per capita energy consumption estimates (using family size as divisor) for the 2747 successfully matched sample families is shown in Figure 1. These are unedited, untruncated values. The range of the FCS values is within rational expectation. On the other hand, the many "outliers" among the FIES values (e.g., > 10,000 kcals) indicate that the data capture method is prone to producing very high measurement errors.

Furthermore, the entire FIES sample, the matched FIES subsample and the FCS sample were subjected to FNRI's range check procedure of rejecting sample points with per capita energy outside the (300, 4000) range. The results showed that 8%, 7% and 0.7% of the samples respectively, fell outside the range (Table 2). Thus, Figure 1 and Table 2 together reveal the clear superiority of FCS over FIES in measuring dietary energy consumption.

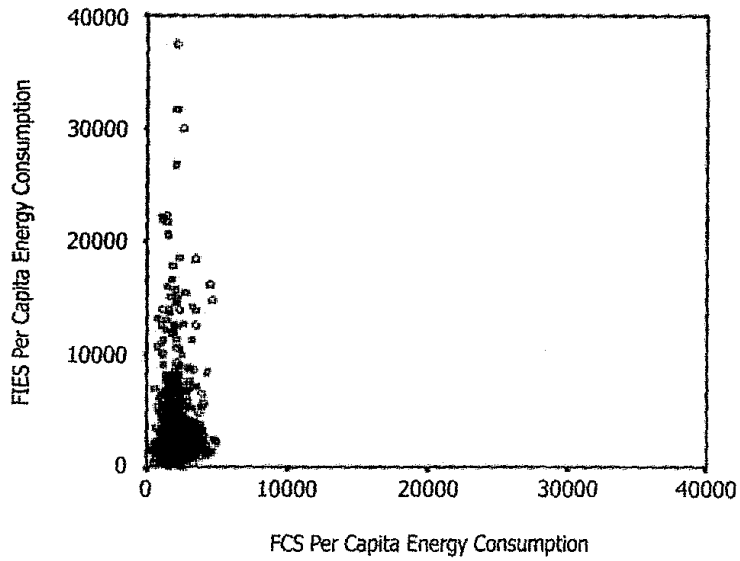
Table 2. **Results of the Range Check Edit Process**

DATA SET	NO. OF OBS	OBSERVATIONS OUTSIDE (300-4000 kcal) RANGE	
		COUNT	PERCENTAGE
FIES	42,094	3,460	8.0
FIES MATCHED WITH FCS	2,747	191	7.0
FCS	3,044	21	0.7

Source: Authors' computations using Family Income and Expenditures Survey and Food Consumption Survey (2003).

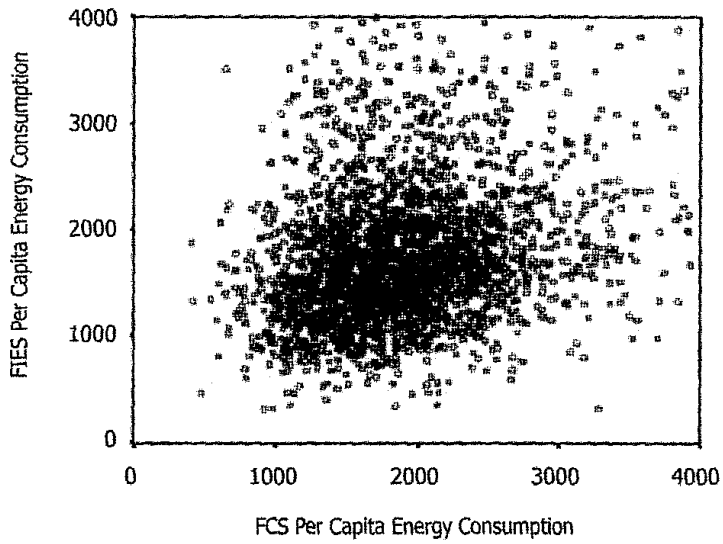
After thorough editing to spot and correct or reject erroneous data points in the FCS sample, FNRI still applied a range check to the remaining subset: per capita kcal values outside (300, 4000) were rejected. Applying the same range check to the FIES-FCS matched dataset left 2541 families whose scatter plot is shown in Figure 2. The correlation coefficient is a *weak 0.20.* This result nipped in the bud our earlier idea to explore combining the two data capture methods in a two-phase sampling scheme where the cheaper interview method is used in a large sample, the more accurate but costlier food weighing method is applied to a much smaller subsample, and either a ratio-type or regression-type formula is used as estimator. It is known, however, that such estimators provide gains in precision only when the correlation between the observations from the two data capture methods exceed 0.5 (assuming equal coefficients of variation; see e.g., Cochran 1977).

Figure 1 **Scatter Plot of FCS and FIES Per Capita Consumption Including Outliers (n=2,747)**



Source: Family Income and Expenditures Survey, Food Consumption Survey (2003).

Figure 2 **Scatter Plot of FCS and FIES Per Capita Consumption within the Range of 300 to 4000 kcal (n=2,541)**



Source: Family Income and Expenditures Survey, Food Consumption Survey (2003).

These findings lead to the recommendation that FIES food consumption data, particularly estimates of energy consumption, be approached with caution when used in hunger or nutrition studies and policy formulation. Pending further empirical investigations, it may be prudent to treat with caution also the other nutrients' estimates similarly derived from FIES quantities of food items consumed.

Furthermore, the above findings lead naturally to questions about the accuracy of estimates of total food expenditure and, by extension, total expenditure which is one of the main reasons for being of FIES. It is an important issue outside the scope of this paper but which, along with estimates of income, we recommend as high priority research [7].

There is little doubt that food weighing as done in FCS leads to very small measurement errors. Questions had been raised, however, whether one day weighing is adequate and whether there can be enough samples to spread in one year to account for seasonality effects. Empirical research comparing the method with more than one day food weighing, daily food recall over a number of days, and combinations of these is highly recommended.

2.5 Recommended Energy and Nutrient Intakes (RENIs) and Adequacy Measures.

Another key table developed and regularly updated by FNRI is that of RENIs for different age-sex groups and for pregnant and lactating women (Table 3). These are also known as, or used to be called, recommended daily allowances (RDAs) for energy, protein, and essential vitamins and minerals.

Averaging each column in Table 3, by using as weights the relative sizes of the rows in the Philippine population from the most recent census, leads to individual RENIs for Filipinos; these are shown in Column 3 of Table 4. The design-based averages from the 2003 FCS sample are shown in Column 2. A comparison of the two columns lead to what FNRI calls percent adequacies in the individual nutrients (Column 4). These show almost 100 percent adequacies for energy and protein, but low adequacy for iron, riboflavin and ascorbic acid. It is to be noted, however, that being based on averages, these do not say much about the number suffering from nutrient inadequacy nor of the severity of the condition. Alternative methods that do not have these limitations are presented in Section 4.

3. The NSCB Official Method

The NSCB is the compiler of the Philippines' official statistics on poverty, including food poverty or core poverty in NSCB's terminology (see e.g., NSCB website). As will be seen shortly, the NSCB method does not use directly either FIES or FCS family food consumption data.

7 Participants in a series of recent United Nations poverty workshops expressed the opinion that food expenditure can be reasonably accurately collected from families; it is quantity that is more problematic, especially when the commodity is not traded in standard units of measure. There was almost full unanimity in this opinion at the 2004 Sub-Regional Workshop of West African States wherein most of the participants were heads of national statistical offices (David 2005).

Table 3 **Recommended Energy and Nutrient Intakes (RENI), Philippines 2002 Edition**

Population Group	Weight *Kg*	Energy *Kcal*	Protein *g*	Vitamin A *µg RE*	Vitamin C *mg*	Thiamin *mg*	Riboflavin *mg*	Niacin *mg*	Folate *µg DFE*	Calcium *mg*	Iron *mg*	Iodine *µg*
Infants, mo												
Birth, < 6	6	560	9	375	30	0.2	0.3	1.5	65	200	0.38	90
6 – < 12	9	720	14	400	30	0.4	0.4	5.0	80	400	10.00	90
Children, y												
1 – 3	13	1070	28	400	30	0.5	0.5	6.0	160	500	8.00	90
4 – 6	19	1410	38	400	30	0.6	0.6	7.0	200	550	9.00	90
7 - 9	24	1600	43	400	35	0.7	0.7	9.0	300	700	11.00	120
Males, y												
10 – 12	34	2140	54	400	45	0.9	1.0	12.0	400	1000	13.00	120
13 – 15	50	2800	71	550	65	1.2	1.3	16.0	400	1000	20.00	150
16 – 18	58	2840	73	600	75	1.4	1.5	16.0	400	1000	14.00	150
19 – 29	59	2490	67	550	75	1.2	1.3	16.0	400	750	12.00	150
30 – 49	59	2420	67	550	75	1.2	1.3	16.0	400	750	12.00	150
50 – 64	59	2170	67	550	75	1.2	1.3	16.0	400	750	12.00	150
65 +	59	1890	67	550	75	1.2	1.3	16.0	400	800	12.00	150
Females, y												
10 – 12	35	1920	49	400	45	0.9	0.9	12.0	400	1000	19.00	120
13 – 15	49	2250	63	450	65	1.0	1.0	14.0	400	1000	21.00	150
16 – 18	50	2050	59	450	70	1.1	1.1	14.0	400	1000	27.00	150
19 – 29	51	1860	58	500	70	1.1	1.1	14.0	400	750	27.00	150
30 – 49	51	1810	58	500	70	1.1	1.1	14.0	400	750	27.00	150
50 – 64	51	1620	58	500	70	1.1	1.1	14.0	400	800	27.00	150
65 +	51	1410	58	500	70	1.1	1.1	14.0	400	800	10.00	150
Pregnant women												
Trimester												
First			66	800	80	1.4	1.7	18.0	600	800	27.0	200
Second		+300	66	800	80	1.4	1.7	18.0	600	800	34.0	200
Third		+300	66	800	80	1.4	1.7	18.0	600	800	38.0	200
Lactating women												
1st 6 mo		+500	81	900	105	1.5	1.7	17.0	500	750	27.0	200
2nd 6 mo		+500	81	900	105	1.5	1.7	17.0	500	750	30.0	200

Note: *Kg* - kilograms; *Kcal* - kilo calories; *g* - grams; *µg RE* - microgram Retinol Equivalent; *µg DFE* - microgram Dietary Folate Equivalent; *mg* - milligrams; *µg* - micrograms

Source: Food and Nutrition Research Institute.

Table 4. **Mean 1-day Per Capita Energy and Nutrient Intake and Percent Adequacy, Philippines, 2003**

Nutrient	Consumption (FCS)	RENI	% Adequacy
Energy (kcal)	1,905.00	1939.00	98.3
Protein (g)	56.20	56.60	99.2
Iron (g)	10.10	16.70	60.1
Retinol (μg)	455.20	498.00	91.4
Thiamin (mg)	0.88	1.02	86.3
Riboflavin (mg)	0.73	1.07	68.0
Niacin (mg)	20.60	13.20	156.4
Ascorbic acid (mg)	46.50	62.00	75.0

Notes: Kcal - kilo calories; g - grams; μg - micrograms; mg - milligrams
Source: National Statistical Coordination Board.

For 2003, NSCB constructed a *per capita one-day menu* for each province that provides 100 per cent adequacy for energy and protein intakes and 80% of the other nutrients including vitamins. The baseline intakes, or *thresholds*, are guided by FNRI's recommended energy and protein intakes (RENIs) seen in Column 3, Table 3. For example, the daily per capita energy and protein intake thresholds are 2,000 kcals and 56 grams respectively, more or less. We say more or less because it will be difficult to construct menus that provide exact threshold specifications. The ingredients used to form the menus were guided by the consumption patterns of "poor" sample families in FCS and possibly FIES. The proposed menus were presented to local field staff for some kind of "validation"; this, however, is not in a statistical sense where the menus would be compared against actual observed or measured consumption by poor families.

Examples of NSCB's per capita one-day menus for Ifugao and Zamboanga del Norte are reproduced in Annexes 1 and 2. The choices are deliberate, with Ifugao on the northern tip and Zamboanga del Norte on the southernmost tip of the country. Ifugao is landlocked, while Zamboanga del Norte has long coastlines in Sulu and Mindanao Seas. Nevertheless, the prescribed one-day menus are identical, except on two counts: for lunch, 5 grams more kangkong (water cress) in Ifugao (90 g) than in Zamboanga del Norte (85 g); for snack 30 gr of pandesal in Ifugao, but 150 gr of saba (boiled banana) in Zamboanga del Norte. The weights of the different ingredients will have to be sorted out in such a way that their total energy and nutrient equivalents lead to the pre-specified values, more or less. Both menus give slightly more (102%) than the prescribed 100% or (2,000 kcals) energy threshold; however, the protein levels (133% and 129%) are much higher than the prescribed 100% or 56 grams. Obviously, the task of finding a menu that will achieve all the prescribed thresholds is much more difficult than it sounds.[8]

8 This fits very well linear programming algorithms which have long been applied to feed formulations for poultry where, given the prices and nutrient contents of the different available ingredients, the problem is to find the cheapest mix that will provide the desired threshold nutrients. There is one big difference: the formulation is actually fed to chickens!

Province level unit prices of the menu's ingredients are obtained by NSCB from various sources, including NSO and BAS. The cost of the menu,

$$\text{fpl} = \sum (\text{weight of ingredients x unit prices})$$

where the summation is through all the ingredients, is NSCB's fpl for the province. Fpls for regions and country are obtained by applying family or population –based weights to compute weighted means.

It may be noted that since menus are not changed frequently, then hypothetically fpl can be computed for any day, week, year or any period that relevant unit prices are available. The reason poverty counts and rates are computed once in three years is that fpl has to be compared with a metric, *per capita family income*, which is updated in FIES once in three years only. That is, a sample family in province A is labeled *food poor or core poor* if its estimated per capita income < fpl in province A. This is the point in the NSCB method where FIES (per capita income) estimates are used directly.[9]

The NSCB 2003 official national level estimates are shown on the upper half of Table 5. The proportions of food poor families and persons are 10.2% and 13.5% respectively, with relative errors (coefficient of variation) 2.2% - 2.3%. Such low CVs are not surprising for estimates of proportions from a large sample of 42,000 families. The NSCB method was applied to the smaller FIES-FCS matched sample families that also have complete per capita income data, with results shown on the lower half of the table. The rise in the CVs is within expectations from the reduced sample size (2,754), and the differences between the incidence estimates and those from the full sample are well within allowable sampling errors. These results validate that there is nothing unusual about the matched sample, which will now be used to examine empirically the performance of the NSCB official method. [10]

Specifically, we look at questions raised on a number of previous occasions (e.g., David and Maligalig 2002). Regarding accuracy, it is reasonable to ask whether or not a one-day menu could represent adequately what poor families eat all year round. No one eats the same menu every day of the year, of course. Even if we assume that the menus (their formulation includes quantities) provide the prescribed RENIs accurately, whether or not the kind

9 Per capita income in the NSCB method is derived as total family income / family size. Results could be significantly different if an adult-equivalent or scale economy of need adjustment is made to family size. An insightful example is the US, where the current poverty line is $10,000 annually for an individual and $20,000 for a family of four (not $40,000); Reuters, in Yahoo News, 29 August 2006.

10 Regional breakdowns are available. However, we limit the analysis here to national level estimates; low level disaggregation will not add value to what we propose to do.

and quantities actually eaten and the kind and quantities prescribed are in close agreement is still an issue.[11] Further, whether the cost of the one day menu multiplied by 365 will come close to the total annual food budget of the poor Filipino family or individual is yet another issue.

Table 5 **National 2003 fpls and Food Poverty Incidence Rates from the NSCB Method**

From Full FIES (n = 42,000)	
Annual Food Poverty Line	8,149 pesos
Food Poverty Incidence	10.2% of families
CV	2.30%
	13.5% of population
CV	2.20%
From FCS (n = 2,754)	
Food Poverty Incidence	9.6% of families
CV	6.20%
	12.0% of population
CV	6.50%

Sources: NSCB website, Arcilla (2006).

The proof of the pudding is in the eating, so the saying goes. We put the 2288 FIES-FCS matched sample families with complete income data through the NSCB method; that is, classify each family as food poor or not by comparing its per capita income with the provincial fpl where the family lives. The results for each province are shown in Annex 3 and summarized in the rows of Table 6. To determine whether the NSCB method has correctly classified the family, the latter's per capita energy consumption is computed from the FCS data and noted if < 2000 (food poor) or not. The results are also in Annex 3 and summarized in the columns of Table 6.

The official NSCB method was able to classify correctly 197 only of the 1,401 families with estimated per capita energy consumption < 2000 kcals. This is a very low 0.14 batting average. It performed better in identifying 842 not food poor from the 887 with per capita consumption ≥ 2000 kcals. The method's empirical success rate is $(197 + 842)/2288 = 0.45$. This is no better than tossing a coin to decide if a family is to be labeled food poor or not food poor based on a 2000 per capita energy consumption threshold.

What do we make of the official 10.2% food poverty rate in 2003 (Table 5)? This is uncanningly close to the marginal frequency $242/2288 = 10.6\%$ in Table 6, which could very well be estimating the proportion of families with per capita incomes $<$ 8,149 pesos. In other words, the official NSCB method divides families more according to per capita income and less on the intended menu-based 2,000 per capita energy consumption threshold.

11 For example, one is entitled to ask why saltwater fish species (galunggong) which would not be readily available in inland mountainous areas, is in the Ifugao menu (Annex 1).

Table 6 **Empirical Validation of Official Method Using FIES-FCS Sample (n=2,288)**

Official Classification	FCS Per Capita <2000 kcal	Energy Consumption ≥ 2000 kcal	Total
Food Poor	197	45	242
Not Food Poor	1204	842	2046
Total	1401	887	2288

Source: Authors' computations using Family Income and Expenditure Survey and Food Consumption Survey (2003).

Previously, the NSCB method prescribed two menus per region, namely one each for urban and rural areas. In 2003 a separate menu had been prescribed for each province, as mentioned above. Presently, the method prescribes separate menus for the urban and rural areas in each province. The remaining steps in this direction would be towns, towns urban and rural, then finally family. The empirical evidence presented above suggests strongly that this direction leads to a dead end.

4. Alternative Methods: Cost of Basic Needs Approach

The discussion here will be confined to methods that use food consumption data directly. Recall that in the NSCB method the food consumption data served only as guide in constructing a one day food menu for each province, and it is the latter that is used to estimate a food poverty line fpl.

Majority of the developing countries that compile poverty statistics follow this approach or some variation of it (David 2005). In this approach, everyone's basic needs may be thought of as falling into two kinds: food and non-food. This paper deals with food needs only.

4.1 Computing fpl through a food basket

Computing food poverty measures involves (i) specifying a food threshold and (ii) determining a *food basket* that satisfies the threshold. For income or expenditure based food poverty measures, (iii) the cost of buying the food basket is fpl; and finally, (iv) a per capita income (or expenditure) distribution estimate is required to determine how many or what proportion of families earn (spend) less than fpl.

As a simplifying assumption, most countries use dietary energy as a proxy for overall nutritional status; and majority of the developing countries use a 2,100 kcal threshold. NSCB's use of 2,000 kcals has empirical basis (see e.g., Column 3 of Table 4 above; also David 2002).[12] The basic data required by

12 As mentioned in sub-section 5.1, the Food and Agriculture Organization has reduced its threshold from close to 2,000 earlier to 1,885 kcals from 2003.

this method are itemized daily quantities of the food consumed by the family which are converted into nutrient equivalents in the manner described in subsection 2.4.

A food basket is constructed by first defining a reference poor population, e.g., some lower percentile of households according to their per capita income or expenditure distribution. The reference population should roughly coincide with, or should contain as subset the poor segment of the population according to most recent estimates. The per-capita food items consumed by this reference population are listed in order of importance, such as with respect to quantity, value, or in some cases frequency of reported consumption by the households. The food basket is comprised of the top entries in this list, stopping at the item where $\sum$ kcal = T' near the prescribed threshold (2,000 for the Philippines).[13]

For time series comparability of the poverty measures it is advisable to keep the same food basket for an extended period. However, changes in consumption pattern could render the food basket inappropriate. Take instant noodles. These have increasingly taken up more space in stores and supermarket shelves. The growing share of instant noodles in the Filipino diet is borne out by FNRI hard data and the agency has begun looking into the food group's effect on the nutrition especially of children. From the 1993 FCS and 2003 FCS, per capita daily noodles consumption has increased by 44%, from 9 grams to 13 grams. However, it may be noted that the NSCB menus do not have noodles. It may be noted further that changes in food consumption patterns affect fpls through changes in nutrient contents and prices of the exchanged items. There is, therefore, some appeal in approaches that circumvent food baskets, like the variation discussed below.

4.2. Eschewing Construction of a Food Basket

A variation that avoids construction of a food basket simply calculates the total food expenditure and total energy consumption of the sample or of a subset representing the reference poor population. The ratio is an estimate of the cost per kilocalorie consumed. For any choice of energy threshold, T, the product

$$\text{fpl} = \text{T} \times (\text{cost per kilocalorie})$$

may be used as a food poverty line. Families or individuals with per capita income or expenditure less than this fpl may be classified food poor. Unlike

13 Since T' $\neq$ 2,000 in general, the sum is forced to the threshold by multiplying each food item's weight consumed per capita by 2,000/T'.

the NSCB method, this simple *cost per kilocalorie method* is computed directly from the foods consumed and prices paid for them by the sample families. Results using 2003 FIES and FCS data are shown in Table 7.

Table 7. **Prices Per kcal and fpls for 2,000 kcal Threshold**

	FIES[a]	FCS-ALL[b]	FCS-LOW33[c]
Mean per capita kcal intake[d]	1739	1887	1494
Per capita food expenditure (¢)	2854	3345	1576
Mean daily price per kcal (¢)	1.706	1.763	1.107
Cost of 2000 kcal			
/Daily (PhP)	34.11	35.25	22.13
/Yearly (PhP)	12,453	12,868	8,080

Notes: [a] n = 2151 families with per capita energy within (300-4000)
[b] n = 2728 families with per capita energy within (300-4000)
[c] n = 838 lowest 1/3 of families in per capita food expenditure
[d] These are unweighted means, which explains the difference with weighted mean in Table 4.

Sources: Authors' computations using Family Income and Expenditure Survey and Food Consumption Survey (2003).

The mean per capita energy consumption in both the FIES (1,739) and FCS (1,887) samples failed to reach the official threshold of 2,000 kcals. Nevertheless, the annual fpls in both cases were above Php12,000 and much higher than the official Php 8,149.

It may be argued that the cost per kcal from the full matched samples is not indicative of the actual cost paid by poor families because the former include non-poor families with dearer food choices and perhaps higher per capita energy intake. The entries in the last column of Table 7 are from the subset of families in the lowest one third in per capita food expenditure. The average cost per kilocalorie for this subgroup was Php. The annual fpl for a 2,000 kcal threshold was Php 8,080 per year, which is roughly two-third of the fpl from the full samples. Thus, the choice of reference (poor) population matters. While Php 8,080 is close to the official Php 8,149, it is to be noted that the former is from families with daily consumptions averaging 1,494 kcals only, while the latter is based on a daily menu that is assumed to provide 2,000 kcals.

Thus, assuming that the FCS per capita energy consumption data are relatively accurate, here are ample empirical evidences that point to the official method's seriously underestimating the number of families that live on less than 2,000 kcals of dietary energy per day.

5. Alternative Methods: Direct use of Consumption, or Non-Money Measures

The CBN and similar methods that involve constructing fpl and comparing these against per capita income (or expenditure) require unit cost data of the different food items and, obviously, family income. Daunting measurement error problems that accompany both data types can impact negatively on the accuracy of the derived food poverty estimates. As mentioned previously (see footnote 7), families tend to know their total expenditure on certain food items better than the amounts consumed and their unit costs. The last of these is usually derived as a ratio of total expenditure/total volume consumed (bought plus own-produced). This is problematic particularly for items purchased in units other than standard weights (e.g., piece, bunch or indigenous measurement units). Family income (expenditure) is notoriously very difficult to collect accurately, especially through face-to-face interview. These practical difficulties make food poverty measures not dependent on prices and income very appealing.

5.1. FAO's Indicator of Dietary Energy Inadequacy

One such measure is the Food and Agriculture Organization's (FAO) proportion of undernourished persons, measured by the *proportion below a minimum level of dietary energy consumption.* FAO's most recent cut-off point is 1,885 kcals per person per day, arrived at in a very complicated manner, by taking into account the amount of food available per person nationally (usually from food balance sheets), a model that represents the extent of inequality in access to food, population growth, and applying an age by sex weighted estimation procedure that considers even the elevated nutritional requirements of lactating and pregnant women (Naiken 2002, FAO Website). FAO's indicator is one of five designated to monitor the first of the UN Millennium Development Goals (MDGs) namely, to reduce extreme poverty and hunger. More specifically, it is meant to monitor progress on MDG Target 2, which is to halve, between 1990 and 2015, the proportion of people who suffer from hunger (UN MDG website). FAO uses it to monitor and report on the world's state of food insecurity and hunger (FAO, SOFI 2003 and later issues). Thus, although the indicator is based on energy consumption only and is aptly described to measure the proportion of persons with dietary energy inadequacy, it has been associated with measuring hunger. A pragmatic way of doing this might be taking progressively smaller cut-off points, and correspondingly associating the resulting indicator values from moderate dietary inadequacy or undernourishment, to malnourishment, or persistent hunger.

The use of dietary energy alone as proxy to overall level of nourishment is based on the assumption that a person whose diet provides him/her adequate energy more likely gets enough of the other nutrients also and vice versa, deficiency in energy intake more likely implies deficiency in some of the other nutrients.

Because FAO compiles its indicator globally with as many countries covered as possible, the agency is compelled to use whatever relevant data is available, including food balance sheets (fbs) in many cases. FBS data, however, is aggregate food supply for human consumption, with no indication of how it is distributed geographically inside the country or among families. Thus many strong assumptions are used to distribute the food supply differentially to families and finally arrive at an estimate of proportion of the population that consumed less than a pre-specified energy threshold. To eliminate these assumptions and to align the FAO indicator with the other indicators (used to monitor MDG goal 1 for instance), the agency needs to revise its methodology, to accept inputs from family/household food consumption surveys instead of national food supply estimates. It appears that FAO is seriously moving in this direction, as mentioned in FAO-SOFI (State of Food Insecurity) and FIVIMS (Food Insecurity and Vulnerability Information and Mapping System) websites; the agency also has started pilot projects using food consumption data have been started in a number of countries.

5.2 Empirical CDF of Nutrient Consumption

Government statisticians routinely construct frequency and cumulative frequency distribution tables from survey data. Constructing tables using per capita energy consumption thresholds or cut-offs as end-points for the classes is just as routine - for self-weighing samples. For unequal probability samples, the appropriate procedure is that of estimating a design-based or weighted empirical cumulative distribution function (empirical CDF).

Let $\Delta(a_i)=\begin{cases}1 & \text{if } a_i \geq 0\\ 0 & \text{if } a_i < 0\end{cases}$ and π_i ; i = 1, 2, ..., n be the inclusion probabilities of the sample families, which in practice are adjusted for non-response, non-coverage and other perturbations in the implementation of the survey. Let x_i be the per capita energy consumption estimate of the i^{th} sample family. The Horvitz-Thompson estimator of the empirical CDF of x is given by (Chambers and Dunstan 1986):

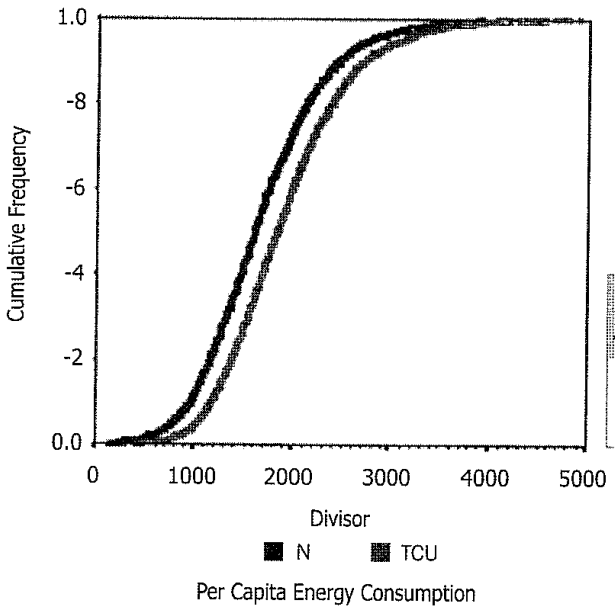
$$\hat{F}(t)=\frac{\sum_{i\in n}\pi_i^{-1}\Delta(t-x_i)}{\sum_{i\in n}\pi_i^{-1}}$$

The procedure is distribution-free or non-parametric because the estimation requires the x values only and not the functional form of F or any of its parameters. Whenever possible, parsimony of assumptions and data requirements are good traits to look for in a statistical procedure. Moreover, the estimator extends readily to more than one variable, such as estimating the joint empirical CDF of per capita energy and protein consumption, which is akin to constructing a two-way table.

In practice, the value of the empirical CDF is calculated for selected points, e.g., deciles or thresholds, and the distribution tabulated or drawn by connecting the points. The results for the FCS sample are summarized in Figure 3 and for two ways of computing per capita values. The left curve is the empirical CDF with per capita energy consumption computed using family size (N) as divisor of total family consumption. The food poverty incidence estimates of 74%, 57% and 35% for 2000, 1700 and 1400 kcal thresholds respectively, are very high compared to the official 10.2% (see table 5).[14] Recall that the official estimate is based on 2000 kcal menus; hence the comparison should be with the 74% CDF estimate.

The right curve replaces N with FNRI's total consumption unit (TCU) that takes into account the family's meal pattern. For example, for a family that regularly eats three meals a day, assign a weight 1, 2/3, 1/3 and 0 to a member that eats three, two one or no meal at home during the food

Figure 3 **Cumulative Distribution of Per Capita Energy Consumption (Kcal) by Different Divisors**



PER CAPITA FOOD POVERTY THRESHOLD	DIVISOR	
	N	TCU
2,000	73.8	60.9
1,700	56.9	41.6
1,400	35.3	21.4

Source: Authors' computations using Food Consumption Survey (2003).

14 Using the matched FIES data, the results, 67%, 51% and 30% are only marginally lower, but also quite high.

weighting day; TCU is the sum of the weights for all members, hence, $0 \leq TCU \leq N$. For a family with a two-meal pattern, the possible weights would be 1, ½ and 0. The resulting food poverty incidence estimates of 61%, 42% and 21% are considerably lower. However, 61% corresponding to a 2000 kcal threshold is still six times higher than the official 10.2% official food poverty incidence estimate. Considering that the design-based mean per capita energy consumption of 1905 kcal (Table 4) is below 2,000 kcal, these empirical CDF results should not be surprising at all.

With the methods described in Maligalig (2007), the incidence of hunger, moderate hunger, and severe hunger are associated with the proportions of respondents that said they missed at least one meal, two meals, and very often missed a meal, respectively, during the reference period. With consumption-based methods, different scales of hunger prevalence can be derived by lowering the energy threshold. A review of international practices and recommendations could serve as guide in determining the lowered thresholds. In nutritional assessment through anthropometric measurements for example, it is standard practice to choose mean – sd and mean – 2sd as thresholds for underweight and severely underweight (wasted) respectively, or short and stunted, where sd is standard deviation of the mean (or standard error). A 1971 FAO/World Health Organization (WHO) Expert Committee on Energy and Protein Requirements accepted a 15% coefficient of variation of energy requirement between individuals in a population or group with similar demographics; (WHO 1985, p.6) Since a 15% cv implies sd = 15% and 2sd = 30%, this means that, with the official threshold as reference point, 2000 × (1-0.15) = 1700 kcals and 2000 × (1-0.30) = 1400 kcals could be proposed to represent thresholds for moderately undernourished and severely undernourished, respectively, or moderately hungry and severely hungry. These are the reasons for having these two additional thresholds in Figure 3.

These results raise two important concerns. First, the FCS per capita energy consumption values seem lower than conventional expectation, resulting in greater than 50% food poverty incidence estimates at the official 2000 kcal threshold.[15] These are much higher than existing estimates both nationally and globally. Removing distortions brought about by per capita calculations reduced the estimates significantly (see next subsection and Table 8); however, the results were still much higher than previously existing estimates). Are the international recommendations on minimum energy requirements, which were the bases for the 2000 - 2100 kcal thresholds used by the majority of developing countries, on the high side to begin with? As mentioned previously, FAO has reduced its threshold to 1885 kcals. Or

15 Judging from Figure 2 and the means in Table 7, the FIES values are comparatively low also.

does one-day food weighing used by FNRI tend to underestimate actual energy intake? Not so, if compared with FIES's face-to-face interview values. Empirical research is badly needed to compare the performance of the current FCS method with others, such as more than one day food weighing, weighing plus additional days of recall, diary method, and combinations of these data capture methods.

Second, TCU adjusts the family size N downward for the effects of meal patterns, but not for scale economies of need, including age and sex variations in RENIs. Moreover, poor families tend to be bigger, as evidenced by higher poverty incidence rates for populations than for families (see e.g. Table 5). Official per capita consumption and income estimates use N as divisor. These mean that on one side, the per capita consumption estimate of poor families will be underestimated, and on another side, their per capita incomes vis-à-vis their real purchasing needs will be underestimated also. The impact of these on food poverty assessment in particular, and on poverty analysis in general, are hard to predict. Empirical studies on alternative methods of per capita calculations from family based basic data are badly needed.[16]

Alternatively, we could consider using methods that circumvent per capita calculations altogether, like the one proposed below.

5.3. Comparing Total Family Consumption with Corresponding Total RENI

Consider a family of four, with husband and wife in the 30-49 age range, an 11 year old son, and a 5 year old daughter. From Table 3, the family's total recommended daily energy requirement is $\sum$RENI = 7,780 kcals, where the summation is through the four members. Let $\sum$kcal be the estimated total energy consumption of the family from a survey, like FCS. If

$$\sum\text{kcal} < \sum\text{RENI} = 7780$$

it is reasonable to say that the family suffers from dietary energy inadequacy; otherwise not. We can consider lower thresholds, like 0.85 ($\sum$RENI) and 0.70 ($\sum$RENI) that more or less correspond with 1 standard deviation and 2 standard deviations less, and call the resulting incidences moderately and severely dietary energy inadequate, respectively. Others may prefer to replace the more technically collect 'dietary energy inadequate' with food poor or hungry.

16 In the mid-1980s Bantilan and associates fitted a double log function of family food expenditure as a function of income and family size, and arrived at a 0.7 estimate of the family elasticity of income. Arcilla (2006) used this provisionally, i.e., $N^{0.7}$, as divisor for computing per capita energy consumption and arrived at much lower food poverty incidence estimates.

In addition to avoiding per capita calculations, totals are self-weighing in the sense that adjustments for the differential nutritional requirements of individuals belonging to different ages and sexes are built in. The method is obviously non-parametric and requires periodic updating of the RENI table and food consumption data only. If desired, energy gaps similar to poverty gaps in mainstream poverty work can be computed simply as the aggregate shortfall between observed consumption and requirement.

The design-based Horvitz-Thompson estimates from 2003 FCS are shown in Table 8. Differences with per capita based estimates using family size N are substantial, e.g., 74% (Fig 3) versus 56% here. It is to be noted that the 17.3% estimate of severe energy dietary energy inadequacy (or hunger) is still higher than the official food poverty incidence of 10.2%. We know now, however, that the latter's empirical link to energy consumption is weak. Possible relationships between this 'total family consumption and RENI comparison" method and the self-assessed poverty approaches of the Social Weather Stations and FNRI would be interesting subjects for future study.

Table 8 Food Poverty Incidence Based on Family RENI Thresholds, 2003 FCS (n = 3,044)

Family Food Threshold	Prevalence (%)	CV (%)
ΣRENI	56.0	1.9
0.85 * ΣRENI	36.6	2.7
0.70 * ΣRENI	17.3	4.5

Source: Authors' computations using Food Consumption Survey (2003).

6. Main Conclusions and Recommendations

The FCS food weighting method yields more accurate results than FIES's interview method. However, the former is very costly, it can cover a fraction only of the FIES sample, and the food weighting is for one day only. We recommend statistically designed field trials to compare different data capture methods, with the aim to find cost-efficient alternatives. We propose FNRI and NSO collaboration here, with an eye for potential further integration and harmonization of FIES and FCS.

The NSCB official method, with its explicit 2000 daily per capita kcal threshold, correctly classified 197 only of the 1401 FCS sample families that recorded less than 2000 kcalories consumption. Overall, it correctly classified 45% only of the FCS sample. We recommend that the method be replaced by a choice from among the alternatives proposed and tried in this study. However, consideration of money-based methods requiring per

capita income (expenditure) and consumption inputs should be preceded by studies designed to produce evidence to support the choice of better ways to generate per capita statistics from family level information.

The method presented in subsection 5.3 above deserves serious consideration. It is not dependent on any assumption, and needs total family consumption ($\sum$kcal) only from a survey. Thus, efforts at improving the accuracy of the food poverty estimates are clearly and unequivocally focused at improving this single input.

7. References

Arcilla, Rechel G. (2006), "Measuring the Prevalence of Hunger in the Philippines Directly from Food Consumption Data." Ph D Thesis, University of the Philippines at Los Banos.

Bantilan, C. S., Pattugalan, J. M., and de Castro, M. (1992), "Adult equivalent income and expenditure distributions: measures of poverty and inequality in the Philippines, 1985 and 1988." Presented at the Consultative Workshop on the Methodologies for an Integrated Analysis of Poverty and Income Distribution and Their Uses for Policy Formulation, 21-24 July 1992, Quezon City, Philippines.

Chambers, R. L. and Dunstan, R. (1986), "Estimating distribution functions from survey data." *Biometrika*, 73, pp.597-604.

Cochran, W. G. (1977). *Sampling Techniques, 2nd Ed.* John Wiley.

David, I. P. (2002) Discussion of Naiken L., "FAO methodology for estimating the prevalence of undernourishment", in *Proceedings: Measurement and Assessment of Food Deprivation and Undernutrition,* FAO, Rome, 26-28 June 2002, pp.43-47.

_____. Country Practices in Compiling Poverty Statistics. Draft chapter 3 in a UN Statistics Division 1995 project, Handbook on Poverty Statistics. New York.

_____ and Maligalig D. (2002), "Issues in Estimating the Poverty Line." *Asia-Pacific Social Science Review*, Vol 3, No. 3, 2002.

FAO. http://www.faostat@fao.org.

FAO. *State of Food Insecurity in the World* 2003 and later issues. Rome.

FNRI (2003), "Food Consumption Survey Report" March 2005.

Mahalanobis, P. (1946), "Recent Experiments in Statistical Sampling in Statistical Sampling in the Indian Statistical Institute." *Jour. Royal Statistical Society*, 109, pp.325-370.

Maligalig, D. (2007), "Examining the Existing Direct Measures of Hunger in the Philippines." Presented at the PSA 2007 Anniversary Conference and published in this issue.

Naiken, L. (2002), FAO "Methodology for Estimating the Prevalence of Undernourishment, in *Proceedings: Measurement and Assessment of Food Deprivation and Undernutrition*," FAO, Rome.

NSCB (2007), http://www.nscb.gov.ph/poverty/FAQs/default.asp.

NSO (2003), Master Sample Documentation, Manila.

UN. http://www.un.org/milleniumgoals/poverty.shtml.

WHO (1985), *Energy and Protein Requirements.* Report of a Joint FAO/WHO/UNU Ad-hoc Expert Consultation. WHO Technical Report Series, No. 724, Geneva.

Examining the Existing Direct Measures of Hunger in the Philippines

Dalisay S. Maligalig[1]

Abstract

The measures of hunger from household surveys conducted by the Bureau of Agricultural Statistics (BAS), Food and Nutrition Research Institute (FNRI) and Social Weather Stations (SWS) were examined to determine which of these can be effectively used in deriving the prevalence of hunger; in identifying at-risk areas and population groups and in monitoring the success of project interventions. On the basis of these measures, there is evidence that the incidence of hunger has increased in the last two years; that hunger has various components and within the households, varying levels of severity. Mothers are more at risk compared to children. Those households with heads who are either unemployed or unskilled workers; whose mothers' highest educational attainment is elementary school are more likely to be hungry. Geographically, the Autonomous Region of Muslim Mindanao (ARMM), has the highest incidence of hunger. The Radimer/Cornell instrument that FNRI used or a good adaptation of it can render consistent and valid direct measures of hunger and food insecurity. However, more research is needed to adapt the instrument so that it can be applied to all types of households.

Keywords: Direct measures of hunger, Radimer/Cornell instrument, household surveys.

1. Introduction

The 2002 World Food Summit concluded that more than 840 million people in the world were hungry. In the Philippines, hunger is also a daily reality for millions. Hunger weakens our bodies and dulls our intelligence. It can hinder a country's growth and development by weakening its workforce. It is, therefore, essential to understand hunger and mitigate this problem.

Hunger is an embarrassing situation that many of us find difficult to confront; making it a complex phenomenon that is not easy to clearly describe and evaluate – a conclusion shared by many researchers in this field. Amartya Sen (1997) surmised that it is difficult to arrive at a clear definition of hunger because the subject matter is dominated by preconceptions and

1 Senior Statistician, Asian Development Bank, 6 ADB Avenue, Mandaluyong City 1550. E-mail address: dmaligalig@adb.org. The views of the author do not necessarily reflect the views of the Asian Development Bank.

often, by attempts to understand a very complex problem in excessively narrow terms. He ventured that hunger can take many different forms and levels of severity.

Many research papers on measuring hunger do not explicitly define it; others equate it with food insecurity while a few with malnutrition, which may also be perceived as a byproduct of persistent hunger. For purposes of this paper, hunger is the most extreme consequence of food insecurity. Food insecurity is defined as the lack of food security wherein food security is a state in which people, at all times, have physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preference for an active and healthy life.

At the national level, there are four existing measures of hunger – (i) the prevalence of food poor (or subsistence poor) which are official statistics compiled by the National Statistical Coordination Board (NSCB); (ii) the self-rated hunger incidence compiled by the Social Weather Stations (SWS); (iii) hunger incidence compiled by the Bureau of Agricultural Statistics (BAS) and (iv) food security measures compiled by the Food and Nutrition Research Institute (FNRI). Three of the existing measures on hunger, those from SWS, BAS and FNRI, were compiled on the basis of responses of individuals to questions about their experiences about hunger. Hence, these are direct or fundamental measures while the prevalence of subsistence poor is indirect or what some researchers call a derived measure (Webb 2006).

The prevalence of food poor or subsistence poor is defined as the proportion of households that have incomes below the minimum requirement to meet the basic food needs which satisfies the nutritional requirements for economically necessary and socially desirable physical activity. A food threshold is determined using regional one-day menus priced at the provincial level. Menus were determined using low-cost, nutritionally adequate food items satisfying basic food requirements of 2,000 kilocalories, which are 100% adequate for the recommended energy and nutrient intake (RENI) for energy and protein and 80% adequate for the RENI for vitamins, minerals and other nutrients (NSCB 2007). The prevalence of food poor is determined by using the food threshold and the income distribution derived from the Family Income and Expenditure Survey (FIES). This measure is discussed in David, et al. (2007).

The SWS self-rated hunger incidence is defined as the proportion of the population who were hungry because they did not have the means to eat. This measure is computed from the results of the SWS quarterly survey – the Social Weather Survey that has about 1,200 respondents from various parts of the country. The respondents are asked if they experienced hunger in the

past three months and how often was this experience (e.g., only once, a few times, often, always). The sampling design and other details of this survey are presented in Appendix A.

The Survey of Hunger Incidence in the Philippines (SHIP) was conducted by BAS in August 2006 to provide indicators that can refine the government's food security programs. This survey administered the same questions asked in the SWS quarterly survey to more than 13,000 respondents nationwide. The details of this survey are discussed in Appendix B.

FNRI adapted the Radimer/Cornell measures of hunger and food insecurity which is based on a set of 10 questions that are designed to evaluate household food security, adult's hunger and children's hunger (Radimer 1990). This set of questions was incorporated in two surveys that the FNRI conducted – the 2001 Updating of the Nutritional Status of Filipino Children (Molano, et al. 2002) and the 6th National Nutrition Survey (NNS) in 2003 (Molano, et al. 2004) that employed a replicate of the National Statistics Office's master sample (NSO 2003). The details of the FNRI measures of hunger and food insecurity are discussed in Appendix C.

This paper examines the measures of hunger compiled by BAS, FNRI and SWS to determine which of these can be effectively used in measuring hunger; in identifying at-risk areas and population groups and in monitoring the success of project interventions. The evaluation is performed on the basis of these types of use and statistical considerations.

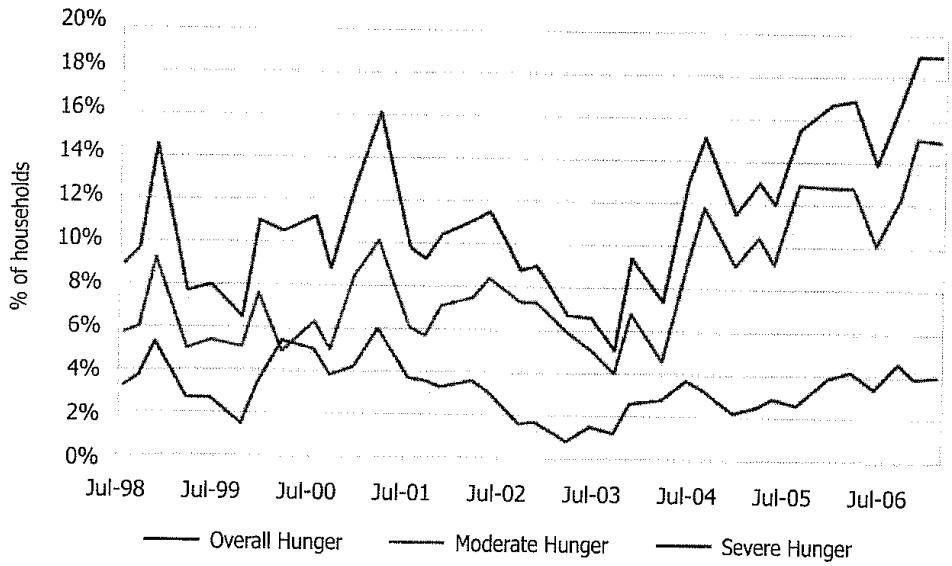
2. How many are hungry?

BAS reports that in the second quarter of 2006, 18.6% of the households in the country experienced hunger; with 3.6% households severely hungry and 15% households hungry a few times. In the same time period, SWS observed that 13.9% of the households experienced hunger, 3.8% of which experienced severe hunger while 10.1% experienced moderate hunger. In fact, per the SWS quarterly survey as shown in Figure 1, except for a downward blip in the second quarter of 2006, the hunger incidence has increased steadily from 12.0% in May 2005 to 19.0% in February 2007.

These figures, however, are moderate compared to the FNRI findings that 84.4% of the households with 0 to 10 year-old children experienced food insecurity in 2001 and 77.0% in 2003. In 2001, FNRI reported that 29.0% of the mothers/caregivers skipped meals so that the children could have a bigger share of the food in the households, while 28.8% of the mothers/caregivers said that they were hungry at least once in six months because

there was no money to buy food. The same proportion of mothers/caregivers (29.0%) skipped meals in 2003, while 24.4% said they went hungry because they had no money. The children were the least affected members of the households with only 15.3% and 15.1% experiencing hunger in 2001 and 2003, respectively.

Figure 1 **Degree of Hunger in Households, Philippines: July 1998 to February 2007**



Source: Social Weather Stations (2007).

The main reasons for the big difference in the FNRI's figures from those of BAS and SWS are: (i) different data collection instruments; (ii) different reference periods; and (iii) different target population. BAS patterned SHIP from the SWS quarterly survey adapting very similar questions on hunger and general sampling strategy (see Appendixes A and B for more details). FNRI, on the other hand, implemented an adaptation of the Radimer/Cornell[2] hunger and food insecurity instrument, which comprised of three components – household, individual/caregiver and children. There were four questions for the first component and three each for the last two components. Each of the questions as indicated in Appendix C is answerable by any of the following: (i) yes (once for the past six months); (ii) yes (at least twice for the past six months); or (iii) never. A household is food secure if the answer to all 10 questions is never. Otherwise, it is food insecure. This very stringent scaling led to the high prevalence of food insecurity. Although the 10 questions are found additive and could be transformed to a single measure, an estimate of the prevalence of hunger was not derived. Moreover, while these two FNRI surveys are of national coverage, the target population for both is limited only to those households with 0 to 10 year-old children.

2 Radimer/Cornell instrument is widely used in the United States and Canada. In Asia, researchers from Malaysia, Pakistan, and Thailand have also used this instrument.

In these two surveys, a household was considered food insecure if it answered positively to any of the four questions for the household food insecurity in the Radimer/Cornell instrument. Unlike other research that used the Radimer/Cornell instrument, a scale for hunger -- which is the most extreme form of food insecurity -- was not established.

As gauge for the reliability of the results, SWS claims the sampling error for the hunger incidence and for all other estimates that came out of their quarterly survey to be about 2.83%. BAS and FNRI did not present any sampling errors in their reports although if warranted, these can be computed from their public use files. The sampling error can represent the reliability of estimates only if these estimates are unbiased. Otherwise, the measure of reliability that has to be considered is the square root of the mean square error (MSE),

$$\begin{aligned} \text{MSE}(\hat{p}) &= \text{E}(\hat{p} - p)^2 + (\text{E}(\hat{p}) - p)^2 \\ &= \text{Var}(\hat{p}) + \text{Bias}(\hat{p})^2, \end{aligned} \tag{1}$$

where $\hat{p}$ is the estimate, p is the corresponding unknown population parameter. The sampling error of the estimate is the square root of $\text{Var}(\hat{p})$, or the variance of $\hat{p}$. The bias of the estimate $\hat{p}$ is nil if the data were taken from a probability sample survey, the correct survey weights have been specified and the responses are reported correctly.

In the case of the SWS quarterly survey, the following design components are potential sources of bias:

- The sampling frame for the first stage of selection was not firmly established. Instead, a sample municipality/city was assigned a number of barangays to be sampled that was roughly proportional to its population size. There was apparently no list of barangays used in the selection. A well-established sampling frame in the first stage of selection must ensure that each primary sampling unit has a chance of being selected, thus eliminating the possibility of bias due to non-coverage.

- Five households were selected from each barangay in the sample. A fixed sampling interval of seven households was used in the selection process. Enumeration stops upon completing five sample households despite the fact that there are other segments of the barangay that have not been covered. This sampling strategy is another version of quota sampling. Since the start of the household count is either the municipal/barangay hall, a school or the barangay captain's house or a church/chapel/mosque, it is likely that the more affluent households will be sampled because in many areas, the poor households are usually located in the periphery of the barangays.

In urban areas, they are in unconventional places like river banks, under bridges, etc. This raises the possibility that hunger incidence could be larger than what is reported by SWS.

While the first-stage sampling frame of BAS SHIP is well-established having used the list of municipalities and barangays from Philippine Standard Geographic Classification, BAS adapted the SWS second-stage sampling procedure and thus, bias is present in the selection of households from the sample barangays.

The survey weights of households in the SWS quarterly surveys do not conform to those of the conventional survey weight components of base weight, adjustment for non-response and adjustment for non-coverage. The household survey weights are uniform within the domains of the SWS quarterly survey and they are equivalent to the number of households that are represented by one sample household. This is usually the base weight which is the inverse of the selection probability. However, if the selection procedure of the SWS quarterly survey is analyzed, the selection probabilities cannot be uniform across domains. Barangays, which are the sampling units in the first stage of selection, vary in terms of the total number of households but they were sampled randomly without considering this variation in the measure of size. Moreover, the sampling interval for the second stage of selection is fixed and does not depend on the total number of households, which are the sampling units in the second stage. Hence, the resulting selection probability, notwithstanding the issue of the non-coverage of some households as mentioned above, will be different across barangays or what SWS calls sample spots. To illustrate, for the National Capital Region (NCR), the probability of drawing a household β in barangay α is as follows:

$$P(\alpha\beta) = P(\beta \mid \alpha\gamma)P(\alpha \mid \gamma) = \frac{5}{M_{\alpha\gamma}} \frac{b_\gamma}{B_\gamma} \quad (2)$$

where $M_{\alpha\gamma}$ is the number of households in barangay α in municipality or city γ, b_γ is the number of barangays to be sampled and B_γ is the total number of barangays in municipality or city γ. The resulting selection probability for each household, therefore, is dependent on the total household size of barangays and the total number of barangays for each municipality or city. This implies that the base weight which is the inverse of the selection probability will vary accordingly contrary to the uniform weights assigned to households in a domain. Without the proper weights, the estimates derived from the survey will not be unbiased and hence, instead of the sampling error, the mean square error should be the basis for the reliability of the estimates (Kish 1992). However, the bias of the estimate cannot be derived

unless a special survey for such purpose is conducted and hence, calculating technically valid measures of reliability for such survey is not possible.

Both BAS and FNRI employed appropriate survey weights in deriving survey estimates but neither computed for the sampling errors of the major estimates. In the case of the FNRI National Nutrition Survey, sampling errors were computed for other characteristics of interest. The sampling errors at the regional level are within the tolerable limit, while those for provinces with small sample sizes are quite large.

For both SWS and BAS surveys, the total sample was allocated equally across domains despite the varying number of households and primary sampling units (PSUs) within domains. There was also no attempt to control for varying base weights within domains. Controls such as further stratification so that the measure of size of PSUs will not vary widely within strata and the use of probability proportional to size selection in the first stage were not considered. Consequently, these factors contribute significantly to the increase in the variance of estimates.

3. Where are the hungry?

In second quarter of 2006, according to BAS SHIP, the top 20 provinces with the highest hunger incidence are presented in Table 1. More than 70% of the households in Lanao del Sur have experienced hunger while the second highest in rank is Siquijor at 51.9%.

Table 1 **Top 20 Provinces with the Highest Hunger Incidence, April-June 2006**

Rank	Province	Hunger Incidence (%)
1	Lanao del Sur	70.8
2	Siquijor	51.9
3	Maguindanao	47.7
4	Biliran	47.1
5	Sultan Kudarat	41.1
6	Pangasinan	37.4
7	Davao Oriental	37.2
8	Davao del Sur	35.4
9	Bohol	35.2
10	Mindoro Oriental	34.9
11	NCR Second District	31.6
12	Western Samar	30.5
13	Zamboanga del Sur	30.4
14	Misamis Occidental	30.3
15	Lanao del Norte	29.0
16	North Cotabato	28.2
17	Nueva Vizcaya	27.9
18	Isabela	27.8
19	Palawan	27.7
20	Negros Oriental	27.5

Source: Bureau of Agricultural Statistics (2006).

Table 2 presents the percentage of households by region who have experienced hunger in 2003 as reported by FNRI and in 2006 as reported by BAS. Although for reasons that have been cited in the previous sections the data of BAS and FNRI are not comparable, it is interesting to note that the ARMM registered the highest percentage in both surveys, 75.2% for FNRI and 35.4% for BAS.

Table 2 FNRI and BAS Regional Estimates of Households who have Experienced Hunger.

Region	Households who were not eating enough, 2003 FNRI (%)	Households who have experienced hunger, 2006 BAS (%)
ARMM	75.2	35.4
SOCCSKSARGEN	68.3	25.1
Western Visayas	66.1	10.9
Eastern Visayas	65.4	20.1
Cagayan	64.5	19.2
Zamboanga Peninsula	61.0	19.1
Bicol	60.0	21.2
Davao	58.9	22.2
Northern Mindanao	58.6	15.1
CARAGA	54.8	13.2
Central Visayas	51.1	17.6
MIMAROPA	49.1	23.6
Ilocos	47.1	23.7
Central Luzon	44.8	9.0
NCR	35.9	24.9
CAR	34.9	5.9
Calabarzon	34.1	9.0

Sources: Molano (2005), Bureau of Agricultural Statistics (2006).

The SWS quarterly survey has 1,200 sampled households and therefore, regional and provincial level estimates with reasonably narrow margins of error are not possible.

To be able to formulate interventions for mitigating hunger and food insecurity, estimates at a finer disaggregation levels are needed. Since the regions do not have governing bodies and funds are allocated to provinces, interventions by the government are usually formulated and implemented at the provincial level. Hence, estimates at the provincial level are more appropriate for purposes of project formulation and monitoring. However, caution must be practiced in interpreting the ranking of these provinces because the ranks may not be distinct especially for those estimates with large sampling errors or mean-squared errors.

To illustrate, Figure 2 presents the percentage of food-poor by province in 2003 (NSCB 2007). The midpoints of the vertical lines are the provincial

estimates of the percentage of food-poor. The upper and lower endpoints of the vertical lines are the upper and lower limits, respectively of the 90% confidence interval[3] of the provincial estimate of the percentage of food-poor. In the figure, Camarines Norte is the 8th highest food-poor province. However, at 90% confidence level, Camarines Norte's proportion of food-poor households ranges from 16.4% to 35%, and therefore, its rank may range between 6 (current rank of Surigao del Norte) to 30 (Abra). Because the ranks are not very distinct, targeting areas that are severely affected and distinguishing them from areas that are at-risk can be difficult. To avoid this situation, the sample-based estimates may be combined with model estimates using small area estimation techniques that could render smaller sampling errors, and consequently, narrower confidence intervals.

4. Who are hungry?

According to BAS SHIP, the households whose heads are laborers and unskilled workers accounted for the highest hunger incidence at 28.7%. Twenty-two percent of those households who experienced hunger had household heads who were unemployed. In ARMM, 40.9% of those households that experienced hunger had agriculture as their main livelihood while 49.0% had household heads who were laborers and unskilled workers.

In 2003, FNRI reported that, 24.4% of the mothers experienced hunger because there was no money to buy more food. On the other hand, only 15.1% of the children experienced hunger. The proportion was similar in 2001 with 28.8% of mothers and 15.3% of children who experienced hunger. Of those who were hungry, 55.8% of the mothers attained elementary education only or none at all. Of those who have experienced food insecurity, 76.5% are unemployed, with 73.1% being housekeepers (Molano 2005).

FNRI's 2001 survey results showed that that 94.1% of the children who were food insecure are underweight (42.3%), stunted (46.7%) and wasted (5.1%).

A profile of households, individuals and children who are at risk of hunger improves the efficiency of the project intervention. Limited resources can be better allocated to those who are at risk and those who are likely to be severely affected by screening them using a well-established profile. The

3 A confidence interval calculated for an estimate of an unknown population parameter, say the mean, shows the range within which the parameter is likely to lie given a fixed error level. Bias must be assessed before confidence intervals can be interpreted because even very large samples that usually result to narrow confidence intervals can be misleading if there are biases in sampling procedures and survey operations.

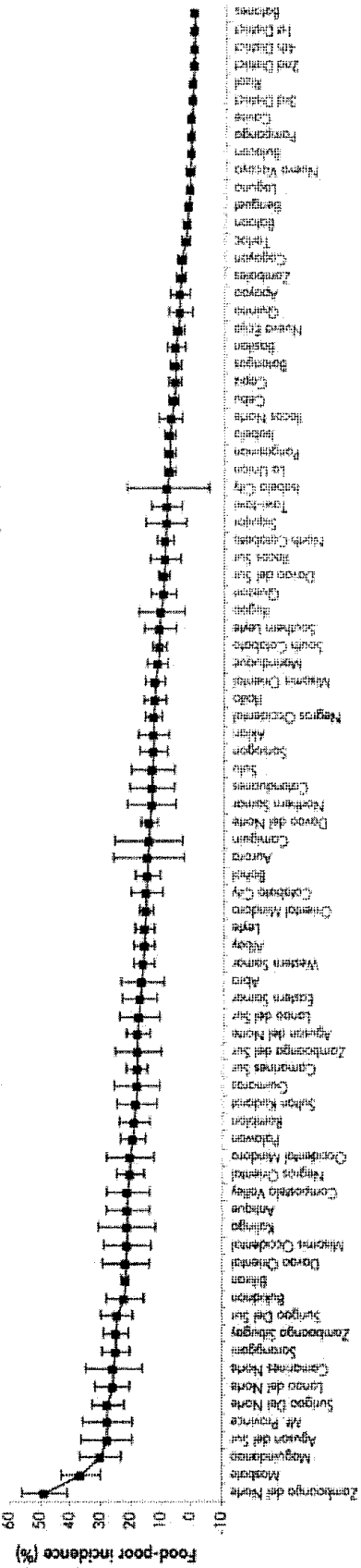


Figure 2. Confidence intervals of provincial estimates of food-poor, 2003.

reports of BAS and FNRI, as discussed above, gave us insights to this profile. Further analysis of their survey data will probably render a better description of who are hungry or at risk. Because the master sample of the National Statistics Office (NSO) was used in the 2003 NNS, its household respondents can be linked with other surveys that were conducted by the NSO; hence, broadening the possible list of determinants of food insecurity and hunger.

It is also important to note that because the Radimer/Cornell instrument that was adapted by FNRI has three components of food insecurity and hunger (i.e., households, adults and children), more in-depth analysis can be undertaken. As mentioned above, it was found that for both surveys that were conducted by FNRI, mothers experienced hunger more than the children. This important result cannot be obtained from either BAS or SWS surveys because their questions are designed such that the response of one member of the household represents all the members of that household.

5. Which measure is the best?

The measures that were examined in this paper are regarded as direct or fundamental measures of hunger, while the official measure that is used for hunger - the prevalence of food or subsistence poor - is considered an indirect or derived measure. Which of these two types of measure is better?

Derived measurement assumes known empirical relationships with the construct of interest, while fundamental measurement does not assume any at all. In general, fundamental measures are preferred over derived measures because of their clear direct link to the characteristic of interest. To put it in another perspective, to ensure the efficacy of a derived indicator on hunger, the relationships between the set of proxy indicators that will be used for the derived measure and the construct of interest that is hunger must first be examined. But then, how can it be examined without first measuring hunger?

Another concern with derived measure is that the relationship of the derived measure with the construct of interest may change in different settings. While there is a direct link between nutritional status and hunger, there are other confounding factors that also affect nutritional status such as availability of health services, level of education of mothers, access to clean water and sanitation. The degree of relationship between nutritional status and hunger may differ in strength because of these confounding factors.

Another way of examining these measures is to classify their data collection methods into objective/quantitative or subjective/qualitative

method. The strength of the quantitative approach is that these methods produce quantifiable, reliable data that could be used to draw conclusions about larger population. This approach fails when the phenomenon under study is difficult to measure or quantify.

An example of the objective-quantitative method is the approach used to derive the prevalence of subsistence or food poor that is based on a monetary line that distinguishes whether nutritional requirements are met or not met by individuals. While this measure provides numerical representation of hunger, the criticism that it is far too distant from the actual experience of hunger has been gaining grounds in recent international research literature (Webb 2006). Within a low-income household, the availability of food or money for food is influenced not only by the household income, but also by food prices, household savings, assets and material support from social networks. At the international level, current research has also raised some issues about the specificity[4] and sensitivity[5] of this measure to individual or household level hunger (Tarasuk 2001).

Subjective/qualitative approach builds on information from the experience and perceptions of individuals. This approach helps researchers understand the meanings individuals assign to social phenomena. Measurement tends to be subjective. The benefit of using the qualitative approach is that it brings in more detailed data from the respondents' perspectives. In many cases, however, the researcher becomes the instrument of data collection and critics can argue that results may vary depending on who conducts the research. Moreover, this approach has not been totally accepted by the mainstream researchers and thus, the validity of the results may be challenged.

The measures that BAS, SWS, and FNRI employed are based on questions regarding the experiences of individuals and households; hence, they can be classified as subjective/qualitative. The common and striking feature of these measures, however, is the transformation of the subjective or qualitative insights from individuals into quantitative measures that can be analyzed using statistical methods. Are these measures valid? There is no gold standard with which to compare these measures; and hence the validity of these measures has to be examined from various perspectives.

In the Philippines, only the FNRI measure was validated. Molano (2006) administered a variation of the Radimer/Cornell instrument (Appendix C) on 3,568 households with children 10 years old or below. These households comprise a subset of the 2003 master sample of the NSO. Cronbach's

4 Proportion of food secure households/individuals that are rated as food secure.

5 Proportion of food insecure households/individuals that are rated as food insecure.

alpha[6] was employed to measure internal consistency while criterion-related validation[7] was conducted by comparing the food security status derived from the adaptation of the Radimer/Cornell instrument with nutritional status and nutrient adequacy. It was found that the prevalence of under-nutrition among children is higher in food-insecure households. Food-secure individuals have higher mean energy and nutrient adequacy. Cronbach's alpha showed agreement between items per food security level in the Radimer/Cornell instrument. On the basis of the results of these analyses, Molano concluded that the Radimer/Cornell instrument was internally consistent and valid in the Philippine setting.

Frongillo, et al. (1996) presented very strong evidence that the Radimer/Cornell and Community Childhood Hunger Identification Project (CCHIP) measures are valid for assessing hunger and food insecurity of households. He assessed the validity of the Radimer/Cornell and similar measures from the CCHIP surveys and the National Health and Nutrition Examination Survey (NHANES) against a definitive criterion measure for hunger and food security in a study of 193 women living with children in a rural county in the state of New York. The criterion measure of food insecurity was derived from independent assessments of the households based on an evaluation of household demographic characteristics, use of food programs, sources and expenditures on food, income, household food inventory, and 24-hour dietary intake recall of the women respondents. The study concluded that there was overall agreement between the CCHIP and the Radimer/Cornell instruments. The Radimer/Cornell and CCHIP measures showed good specificity and excellent sensitivity. Although it displayed excellent specificity, the NHANES measure significantly underestimated the prevalence of food insecurity and resulted in very poor sensitivity. This result does not imply that there is something wrong with NHANES question but rather, it supports the contention that there is no single item alone that is sufficient for assessing hunger and food insecurity.

6 Cronbach's alpha evaluates how well a set of items (or variables) can numerically represent a single unidimensional latent construct. When data have a multidimensional structure, Cronbach's alpha will usually be low. Cronbach's alpha is not a statistical test. It is a coefficient of reliability (or consistency).

7 In psychometrics, criterion validity is a measure of how well one or set of variables/measures predicts an outcome based on information from other variables/measures. Criterion validity is achieved if a set of measures from a personality test relate to a behavioural criterion that experts agree on.

Derrickson (2000) validated the Core Food Security Module (CFSM)[8], which is very similar to the Radimer/Cornell instrument to a sample of 1,603 Hawaii residents, 198 of which were Filipinos. She concluded that CFSM was valid across all ethnic groups that were studied and that her findings were consistent with previous research. Other key researches also support the validity of the Radimer/Cornell, CFSM, CCHIP measures as applied in the United States, Canada and other countries.

Although the validity of the SWS and BAS surveys has not been closely examined, insights from the abovementioned studies can be drawn since the question pertaining to hunger in these two surveys is the same as that of the NHANES; and also, the same as that of item 3 of the original Radimer/ Cornell survey instrument. On the basis of the Frongillo (1996) study, the specificity of the NHANES survey is excellent but its sensitivity is quite poor and therefore, compared to the other measures, the NHANES prevalence of food insecurity at the household level was also significantly lower. This implies that one question may not be sufficient for assessing hunger and food insecurity.

It is interesting to note that some results from the BAS SHIP as presented in Table 3 indicate that hunger incidence may not be consistent with the socio-economic classification. Enumerators, using a very comprehensive classification matrix, designated the socio-economic classification of households. The matrix includes description of the general appearance of the housing unit, the occupation of the head of the family and available household conveniences and appliances. Table 3 shows that in areas in Luzon that are outside NCR, 17.4% of those in the upper class (AB) experienced hunger, much higher than those in middle and lower classes (C and D). In Mindanao, 6.9% of those in the upper class experienced hunger compared to 2.8% in the middle class. These results imply that either the socio-economic classification was not properly designated in these areas or that the specificity of the measure is much lower than expected.

Table 3 **Percentage of Households that Experience Hunger by Socio-economic Classification**

Region	Socio-economic classification			
	AB	C	D	E
NCR	0.0	0.0	13.8	63.0
Balance of Luzon	17.4	1.9	10.4	39.1
Visayas	0.5	0.5	11.9	37.1
Mindanao	6.9	2.8	13.4	47.0
Philippines	7.2	1.3	12.0	48.7

Source: Bureau of Agricultural Statistics (2006).

8 The United States Department of Agriculture developed the Core Food Security Module as a national measure of hunger and food insecurity. This module is a set of 18-item scale that was formulated on the basis of CCHIP and Radimer/Cornell items.

In terms of reliability of results, the estimates at the regional level from the 2003 National Nutrition Survey have tolerable sampling errors. Because of the bias that is introduced in the second-stage selection, the sampling errors of estimates from BAS and SWS surveys may not be appropriate and the mean square errors of the estimates, however, cannot be calculated.

Finally, for lack of data, the comparisons that were made here did not consider the resource requirements for developing, validating and compiling the existing direct measures of hunger. These resource requirements should also be weighed against the available indirect measure.

6. Conclusions

What does current research on hunger tell us? There is evidence that the incidence of hunger has increased in the last two years; and hunger has various components and within the households, varying levels of severity. Mothers are more at risk compared to children. Those households with heads who are either unemployed or unskilled workers, or whose mothers' highest educational attainment is elementary school are more likely to be hungry. ARMM has the highest incidence of hunger.

The measures that were used to draw these conclusions are many and varied. Considering the discussions in the preceding sections, it is clear that there is no perfect single measure that was able to capture all aspects of hunger including its level of severity and address all the requirements for purposes of project/policy formulation and monitoring. Of the three direct measures of hunger, the SWS prevalence of hunger has the most frequent observations (every quarter) and the longest data series, having started in 1998. However, because of its limited sample size, it cannot be used for identifying at-risk areas smaller than the survey's domains. Estimation procedure can still be improved by specification of the correct survey weights and the possible bias due to the non-probability sample can be eliminated with the use of probability sampling design.

Because of its adequate sample size, the SHIP, which was patterned after the SWS quarterly survey, may be used to target at-risk areas. However, it also suffers from the same criticism about the use of non-probability sample. Because of this, sampling errors cannot be computed accurately and thus, the ranking of provinces will not be clearly delineated. SHIP was only conducted once, hence, unless BAS conducts it at regular intervals, it cannot be used for monitoring.

Due to the process of selection of households that was used by BAS and SWS, the estimates of the prevalence of hunger may be biased and probably underestimated. For both the BAS SHIP and SWS Quarterly Survey, the question on hunger that leads to the computation of prevalence may also have lower sensitivity that can result in lower estimates as implied by more advanced research. This conjecture, however, which is based on research results in other countries, should be further investigated in the Philippine setting.

The FNRI measures of hunger and food insecurity were derived using an adaptation of the Radimer/Cornell instrument, which has gained prominence in the United States and Canada. Researchers found the instrument and adaptations of it valid and internally consistent. The FNRI data collection instrument can distinguish different levels of food insecurity at the household level and within the household. It is rather unfortunate, however, that FNRI was not able to apply methods for distinguishing households, individuals and children who are at risk, those who are food insecure from those who are severely hungry. Moreover, the survey instrument was only applied to households with 0-10 year-old children, limiting the coverage and conclusions that can be drawn from the results.

Since FNRI is willing to share its data, further research can be done to address the scaling problem mentioned above and also, to broaden the search for determinants of hunger by linking the 2003 NNS with other surveys conducted by the NSO.

It is also noted that for screening, or for identifying households that are at risk of hunger, excellent sensitivity is more important than specificity. It is better to be able to identify all potential at-risk households and introduce further evaluation measures to separate those who appear to have food problem but do not.

This paper seeks to generate better understanding of these strengths and weaknesses of existing direct measures of hunger so that users are better informed to choose which of them is best suited for their purposes. What has emerged in this search for the best measure of hunger and food insecurity is that a valid direct measure is possible. However, more research is needed to obtain an ideal measure – one that can identify at-risk areas; that can be used for screening households and individuals; that is measured at a regular interval to monitor this very dynamic phenomenon and that is consistent, sensitive and reliable.

7. Recommendations

While we cannot identify which is the best of the existing direct measures of hunger, we also do not recommend that an entirely new measure be developed for such purposes; rather we hope to combine the strengths of existing direct measures of hunger and food insecurity; and learn how to address their perceived weaknesses. The existing direct measures are found to be valid and with further refinements, they could approximate the ideal measure that we seek which could be used for policy and project formulation and monitoring.

SWS showed us that hunger is a very dynamic phenomenon. The current state of hunger cannot be fully represented by estimates from a survey that is conducted once every three years. Hence, to be able to mitigate hunger, it should be measured regularly and at shorter intervals. However, the results of the National Nutrition Survey and the Family Income and Expenditure Survey should not be discarded. These surveys provide strong data support for more in-depth analysis of this phenomenon. Also, their results could be used to validate the direct estimates of the prevalence of hunger.

A survey similar to that of the SWS quarterly survey is the appropriate vehicle for the final instrument that must be developed for monitoring the prevalence of hunger at the national level. With a good sample design and adequate sample size, the results of such survey can also be used for targeting at-risk areas. The sampling design and estimation procedures of such survey should be closely examined to ensure the reliability of results

Given the limited resources for official statistics, an independent survey on hunger and food insecurity at regular intervals may not be possible. If this is the case, then the instrument for direct measures of hunger can be administered as a rider to the quarterly Labor Force Survey (LFS). On the basis of research in other countries, the Radimer/Cornell and similar instruments do not require ample time to administer. Tarasuk (2001) determined that it takes four minutes to administer the CFSM with 18 questions for households with children. A time and motion study in the Philippine setting could validate this result.

Because of the sample design, however, unemployment estimates from the LFS for provinces with small sample sizes have sampling errors larger than the tolerable level. This result may also be true for the prevalence of hunger; and hence, small area estimation techniques must be considered to generate estimates at the provincial level. This area is also well researched in more advanced countries and the availability of literature for extending this type of research in the Philippines will not pose any problem.

To enrich research on how hunger and food insecurity can be abated, surveys on hunger and food insecurity must also capture other characteristics of the households and individuals. If the proposed survey for direct measures of hunger is administered as a rider to LFS, then the survey can be linked to other surveys that also use NSO's master sample since they share a common set of households. And as demonstrated by David, et al. (2007), when the National Nutrition Survey was linked to FIES, the income level of households can be determined vis-à-vis its nutritional status and thus, more conclusions can be drawn. Moreover, the relationships of indirect measures such as prevalence of food-poor and poverty incidence with direct measures of hunger and food insecurity can be examined.

The use of probability sample survey and the correct weights should be promoted. To design the sample survey carefully, the availability of sampling frame, measure of sizes, the contribution of disproportionate selection or varying weights to the sampling error and limited resources must all be considered. And to support the validity of conclusions drawn from surveys, estimates should be published or disseminated with their corresponding sampling errors.

As demonstrated by the FNRI, the Radimer/Cornell instrument or a good adaptation of it can render consistent and valid direct measure of hunger and food insecurity. However, more research is needed to adapt the measurement so that it can be applied to all types of households. The following may be considered:

> Kendall (1995) suggested that the nutritional aspects of the diet (quality) will help delineate those households that are food insecure but not yet experiencing the hunger; thus, she recommended additional items like "I can't afford to buy the foods that I think I should feed my family," may be added in the instrument to assess diet quality in order to accurately estimate the prevalence of individual-level food insecurity.
>
> Frongillo, et al. (1996) recommended that cut-off points in the food insecurity scale (total number of affirmative responses) must be made on the basis of the substantive meaning of the levels of food insecurity and that the scale must have sufficient number of items – at least four and preferably five or more – for measuring each level of food insecurity.
>
> Researches following the groundbreaking research of Radimer, suggest three choices for responses to questions: "never, sometimes, often" and "not true, sometimes true, often true" for statements,

instead of the five choices that were implemented in the original research questionnaire. Also, respondents have indicated that statements were easier to understand than questions.

If the instrument is administered to all types of households – with or without children, how will the cut-off points for hunger and food insecurity be determined? The questions/statements on children will not apply for those households without children and therefore, the use of additive scale has to be modified. In addition, how will the prevalence of hunger be computed with the possibility of difference in hunger status within households for children and adult? Radimer suggests that for monitoring hunger, the frequency distribution by each of the component (household, adult, children) of the scale scores should be studied and cut-off points should be set on the basis of the frequency distribution of scores.

To maintain consistency and ensure that change is being measured accurately, the same instrument for both monitoring and targeting purposes should be employed. Moreover, the same instrument can be used for screening households/individuals at the project level. If this is done, care must be taken in administering the instrument. Enumerators/project staff must be trained to elicit the appropriate answers to the items in the instrument. For example, current research in the United Sates and Canada suggests that the interviewer should use carefully designed follow-up questions and repeat the response and item to verify response.

Appendix A: The Social Weather Survey

1. Introduction

This note is based on the documentation called Third Quarter 2004 Social Weather Survey Technical Details. This survey is done quarterly by the Social Weather Stations (SWS) using a 36-page questionnaire that asks a variety of political and social issues. Most of the questions correspond to indicators that are monitored over time by SWS. The questionnaire was administered to voting-age adults (18 years old and above).

The two questions that pertain to hunger are reproduced below:

E. HUNGER

12. Nitong nakaraang tatlong buwan, nangyari po ba kahit minsan na ang inyong pamilya ay nakaranas ng gutom at wala kayong makain?

    In the last 3 months, did it happen even once that your family experienced hunger and not have anything to eat?

    OO *(Yes)* 1

    HINDI *(No)* .....2 ✍ **GO TO Q14**

13. KUNG OO: Nangyari po ba 'yan ng MINSAN LAMANG, MGA ILANG BESES, MADALAS, o PALAGI? (SHOWCARD)

    IF YES: Did it happen ONLY ONCE, A FEW TIMES, OFTEN, or ALWAYS? (SHOWCARD)

    MINSAN LAMANG (Only once) 1

    MGA ILANG BESES (A few times) 2

    MADALAS (Often) 3

    PALAGI (Always) 4

2. Coverage and sampling frame

The survey covered the entire Philippines. It has four domains – the National Capital Region (NCR), Balance of Luzon (outside NCR), Visayas and Mindanao.

3. Determination of the sample size

The technical documentation does not discuss the basis for determining the total sample size. It seems that the specification that five households will be taken from each barangay that will be selected and that 300 barangays will be selected from each of the domains mentioned above was done without data support from previous surveys' methodological analysis and assumptions on

acceptable error level. There were 240 barangays (or sample spots) that were selected in the survey and 1,200 respondents representing the same number of households.

4. General sampling design

The following paragraphs are quoted verbatim from the Technical Details paper.

For the National Capital Region

Stage 1. Selection of Sample Spots (Barangays)

For NCR's first stage, 60 barangays are distributed among the 17 NCR cities and municipalities in such a way that each city/municipality was assigned a number of barangays that was roughly proportional to its population size. An additional provision was that each municipality must receive at least one barangay. Barangays were then selected at random from within each municipality.

Stage 2. Selection of Sample Households

In each sample barangay, interval sampling is used to draw 5 sample households: A starting point in the spot map was assigned at random. The first sample household was selected randomly from the households nearest to the starting point. Subsequent sample households were chosen using a fixed interval of 6 households in between the sampled ones; i.e., every 7th household was sampled.

Stage 3. Selection of Sample Adult

For the third stage, in each selected household, a respondent is randomly chosen among the household members who were 18 years of age and older, using a probability selection table. In selecting the probability respondent of a household, only male family members were pre-listed in the probability selection table of odd-numbered questionnaires; only female family members were pre-listed for even-numbered questionnaires. In cases where there was no qualified probability respondent of a given gender, the interval sampling of households would continue until five sample respondents were identified.

For the rest of the Philippines

Stage 1. Allocation and Selection of Sample Provinces

Balance Luzon was further divided into 5 regions: CAR + Region II, Region I, Region III, Region IV and Region V; Visayas into 3 regions: Region VI, Region VII and Region VIII; and Mindanao into 6 regions; Region IX, Region X, Caraga, Region XI, Region XII, and ARMM.

Using probability proportional to population size (PPS) of the region, the allocation of 10 provinces in Luzon, and 5 in Visayas and 6 in Mindanao were as follows:

LUZON		VISAYAS		MINDANAO	
CAR/Region II	1	Region VI	2	Region IX	1
Region I	1	Region VII	1	Region X	1
Region III	2	Region VIII	1	Caraga	1
Region IV	3	Non-quota	1	Region XI	1
Region V	1			Region XII	1
Non-quota	2			ARMM	1
TOTAL	10		5		6

The non-quota provinces were selected without replacement using probability proportional to their remainders. The remainders are fractions derived when the proportion of the regions (based on their respective study area) are multiplied by 10 for Luzon, and 5 for Visayas and 6 for Mindanao. For instance, if 1.45 is obtained for Region I, then 1 province is assigned to this region and remaining fraction of 0.45 is included for the allocation of the non-quota province.

Given the quota for each region, sample provinces were then selected by PPS, without replacement. An additional provision is that each region must receive at least one province.

Stage 2. Allocation and Selection of Sample Spots

Once the sample provinces have been selected, 60 spots for each of the major areas were allocated among the sample provinces. Using the quota set for each spot in each region, the spots were distributed in such a way the each province was assigned a number of spots roughly proportional to its population size.

Sample barangays within each sample province were randomly selected with equal probabilities.

LUZON		VISAYAS		MINDANAO	
CAR/Region II	8	Region VI	24	Region IX	10
Region I	8	Region VII	22	Region X	9
Region III	15	Region VIII	14	Caraga	7
Region IV	21			Region XI	17
Region V	8			Region XII	9
				ARMM	8
TOTAL	60		60		60

To yield representative figures at the national level, census-based population weights are applied to the survey data. The weight projection is computed by dividing the projected population in the area by the sample size of the same area. Appropriate projected factors were applied so that original population proportions were reflected in the data tables using this formula.

5. Survey weights

Weighting Procedure

To yield representative figures at the national level, census-based population weights are applied to the survey data. The weight projection is computed by dividing the projected population in the area by the sample size of the same area. Appropriate projected factors were applied so that original population proportions were:

$$\text{Projection factors} = \frac{\text{Population}}{\text{(Weight) No. of Interviews}}$$

	2004 NSO Projected Population Age 18 and above	Total Sample Size (PR)	Projection factor
NCR	7,245,306	300	24.15
Balance Luzon	20,796,326	300	69.32
Visayas	9,777,073	300	32.59
Mindanao	11,304,824	300	37.68
TOTAL	49,123,530	1,200	

For questions answered by the sample voting-age adults, the following projection factors were used:

	2004 NSO Projected Household Population	Total Sample Size	Projection factor
NCR	2,335,161	300	7.78
Balance Luzon	6,889,866	300	22.97
Visayas	3,286,814	300	10.96
Mindanao	3,806,091	300	12.69
TOTAL	16,317,932	1,200	

For questions pertaining to household (HH), the following projection factors were used:

The SPSS version of the datafile is already weighted according to the above projection factors. As the data are weighted, the total number of cases that appear is 49,124. The figure is in thousands, i.e., 49,123,530 persons representing NSO's projected number of adults (18 years old and above) for year 2004 based on the 1995 Census.

Researchers who are defining data using the ASCII files should apply these projection factors.

Note: This Appendix was taken from the Social Weather Stations (2004).

Appendix B: Survey Of Hunger Incidence In The Philippines

1. Introduction

The Survey of Hunger Incidence in the Philippines (SHIP) was conducted to address the requirement of the Department of Agriculture for determining the hunger situation and its underlying causes in the country. At the time of the survey in 2006, the Social Weather Stations (SWS) first quarter survey report indicated a steady percentage of households that were suffering from hunger, 16.7% and 16.9% in the fourth quarter of 2005 and first quarter of 2006, respectively. Food production statistics were also on the uptrend and therefore, causes of hunger other than availability of food should be identified.

2. Coverage and sampling frame

The survey covered 78 provinces, the National Capital Region (NCR) and the chartered cities of Zamboanga and Davao. The list of municipalities and barangays taken from the Philippine Standard Geographic Classification (PSGC) as of 31 March 2006 served as the sampling frame.

The provinces, districts of NCR and the two chartered cities were the domains of the survey.

3. Determination of the sample size

The technical documentation does not discuss the basis for the total sample size. It seems that the specification that 10 households will be drawn for each of the 16 barangays that will be drawn per province was done without data support from other similar surveys and assumptions on acceptable error level. There were 1,291 barangays and 12,857 sample households that were surveyed.

4. General sampling design

In general, the sampling design is similar to that of the SWS -- two-stage with the barangay as the primary sampling unit and households as the ultimate sampling unit. For each province, 16 barangays will be selected, from each of which 10 households will be drawn. On the other hand, for NCR, two barangays will be drawn for each municipality, for each of which 10 households will be selected.

Unlike the SWS quarterly survey, however, some stratification mechanisms were introduced in SHIP. For each province, district of NCR or chartered

city, the barangays were first stratified according to urban/rural classification. For each stratum, the barangays were then sorted according to the percentage of underweight children in a municipality. This implies that the barangays within a rural/urban stratum will be sorted by municipality in the order of percentage of underweight children from the 2004 Operation Timbang of the National Nutrition Council. The barangays are then drawn systematically with the starting number drawn between 1 and the sampling fraction which is $\frac{B_{ij}}{b_j}$, where B_{ij} is the total number of barangays, while b_{ij} is the total number of barangays to be sampled in the jth stratum of the ith province. b_{ij} is derived proportionately such that 16 will be total sampled barangays per province.

Since there is no complete listing of households in the selected barangay, the 10 households per barangay are then selected systematically with that the starting number drawn between 1 and 6, for urban barangays; or between 1 and 3 for rural barangays. The starting point of the count of households is either a barangay hall, school, barangay chairperson's house, or church. The enumerator moves to the right of the starting point, choosing households along the road or passageway. The enumerator will go through the alleys intersecting the main road in a serpentine manner.

The first sampled household will be the household with the number corresponding to the random start. If the selected household refused to participate, or if there is no qualified respondent present in the household then the data collector will proceed to the adjacent household. Then the enumerator will go to the next household to be sampled which is either the sixth (for urban) or the third (for rural) from the current sampled household. This process continues until there are 10 sampled households.

5. Survey Instrument

The survey questionnaire is a two page questionnaire that lists all the members of the household, their ages, sex, highest educational attainment and main occupation; total household income for the first 6 months of 2006, consumption of staple foods; awareness of specific government programs and inventory of palay, rice and corn in the household, in addition to the two major questions about hunger incidence: (i) During April to June 2006, did it happen even once that your household experienced hunger and have nothing to eat? If so, how often; what was the main reason why the household experienced hunger? (ii) During April to June 2006, how many meals and snacks did your household normally eat per day?.

6. Estimation procedure

On the basis of SHIP's technical notes, given that n_{ijk} is the number of households to be selected in the k^{th} barangay in the j^{th} stratum (urban/rural) and the i^{th} province; N_{ijk} is the total number of households, then the selection probability of each household is:

$$P(ijk) = P(k \mid j)P(\textit{choosing a household} \mid k)$$
$$= \frac{b_j}{B_j}\frac{n_{ijk}}{N_{ijk}}$$

Then the base weight for each household, w_{ijk} is:

$$w_{ijk} = \frac{1}{P(ijk)} = \frac{B_j}{b_j}\frac{N_{ijk}}{n_{ijk}}$$

If

$$Y_{ijkl} = \begin{cases} 1 & \textit{if the lth sample household experienced hunger,} \\ 0 & \textit{otherwise,} \end{cases}$$

then the prevalence of hunger at province i is estimated as:

$$\hat{P}_i = \frac{\sum_j \sum_k \sum_l w_{ijk} Y_{ijkl}}{\sum_j \sum_k \sum_l w_{ijk}}$$

The estimation procedures for NCR and the two chartered cities follow the same pattern. There was no formula given for the sampling error. Although, the Taylor series linearization method should be appropriate for this ratio type estimator.

Note: This appendix was reproduced from the Bureau of Agricultural Statistics (2006).

Appendix C: Food Insecurity Measures By The Food And Nutrition Research Institute

1. Survey Instrument

The measures of food insecurity and hunger that FNRI have produced are based on an adaptation of the Radimer/Cornell instrument that was developed by Radimer (1990) based on in-depth interviews with 32 women with children living at home who have experienced hunger. A narrow and a broad concept of hunger emerged from those interviews. The narrow concept equates hunger to the physical sensation of hunger pangs, insufficient food intake and going without food, while the broad concept includes problems with household food supply, quality of diets, feelings about the situation and what was done to maintain household food supplies – which corresponds to food insecurity.

Radimer (1990) and many research on food insecurity and hunger that followed claim that food insecurity is experienced differently at the household level and within the household, the experience differs between adults and children. The table below presents a comparison between the FNRI and the original Radimerl/Cornell instrument.

Comparison Between FNRI and the Original Radimer/Cornell Instruments	
FNRI **Household Food Insecurity**	**Radimer/Cornell** **Household Hunger**
1. I worried that our food would run out before we got money to buy more.	1. Do you worry whether your food will run out before you get money to buy more?
2. The food we bought did not last and we do not have enough money to get more.	2. The food that I bought just didn't last, and I didn't have money to get more.
3. The children were not eating enough because we did not have enough food and we could not afford to buy more.	3. I ran out of the foods that I needed to put together a meal and I didn't have money to get more food.
4. We could not feed the children a nutritionally adequate meal because we do not have enough food and enough money to buy food.	4. I worry about where the next day's food is going to come from.
Adult's Hunger	**Women's hunger**
1. Did you skip eating or miss meals/food, because there was no food or no money to buy food?	1. I can't afford to eat the way I should.
2. Did you ever not eat for a whole day, because there was no food or money to buy food?	2. Can you afford to eat properly?

FNRI Household Food Insecurity	Radimer/Cornell Household Hunger
	3. Do you eat less than you think you should because you don't have enough money for food.
Children's Hunger	**Children's hunger**
1. Did your child/children skip eating or miss meals/food, because there was no food or no money to buy food?	1. I cannot give my child(ren) a balanced meal because I can't afford that.
2. Did your child/children ever not eat for a whole day, because there was no food or money to buy food?	2. I cannot afford to feed my child(ren) the way I think I should
3. Was/were your child/children ever hungry but did not eat because there was no food or money to buy food?	3. My child(ren) is/are not eating enough because I just can't afford enough food.
	4. I know my child(ren) is/are hungry sometimes, but I just can't afford more food.

Sources: Molano (2005), Radimer (1990).

The Radimer/Cornell instrument has been applied in many surveys in the US, Canada and other countries, to provide categorical determination of household food insecurity, individual (adult) food insecurity, child hunger, and individual-level hunger for the adult respondent. The scales are additive and hunger is determined based on the number of affirmative responses.

2. Data Collection Mechanisms

The FNRI adaptation of the Radimer/Cornell instrument was implemented in two surveys – in the 2001 Updating of the Nutritional Status of Filipino Children and in the 2003 National Nutrition Survey.

2001 Updating of the Nutritional Status of Filipino Children

The domains of the survey are the 16 regions. However, the sampling strategy focused on 19 provinces with the highest prevalence of underweight 0-5 year-old children as found in the 1998 Antropometric Survey. While barangays were stratified according to provinces, these 19 provinces were allocated enough sample sizes such that stable estimates can be generated at the provincial level.

A two-stage sampling design was used in which, barangays were selected in the first stage and children were selected in the second stage. The list of barangays for each province was taken from the 2000 Census of Population and Housing. For each barangay selected, listing operations were conducted from which households with 0-10 year old children were drawn.

It is not clear how the number of barangays were allocated to provinces and the total sample size was determined. A total of 599 barangays and 12,425 children were selected for this survey

2003 National Nutrition Survey

The 2003 NNS used one replicate of the National Statistics Office's master sample, which has four independent replicates in total. The domains of the master sample are the regions. In general, the sampling design of the master sample is two-stage in which the primary sampling units (PSUs) can be barangays or for large barangays, enumeration area (smaller than barangays) and housing unit is the ultimate sampling unit. Stratification was also introduced in the first stage of selection. PSUs were stratified by province (including highly urbanized cities and independent component cities), percentage of housing units occupied by households that are made of strong materials, whether majoring of households in the PSU are engaged in agriculture and the per capita income of the municipality in which the psu is located.

Sample size was determined on the basis of a threshold sampling error level of major variables from past surveys of similar coverage and the sample size was allocated across domains using Kish Allocation. The full details of the sampling design, including the specification of weights, rotation and estimation procedures are in the 2003 Master Sample Documentation.

Although NNS was applied to one replicate of the 2003 Master Sample, the instrument on food insecurity and hunger was only administered to households with 0 to 10-year old children, or 3,568 in 786 enumeration areas nationwide.

References

Bureau of Agricultural Statistics (2006), Survey of Hunger Incidence in the Philippines. Unpublished Technical Report.

_____ (2007), Key Findings of the Survey of Hunger Incidence in the Philippines. Unpublished Technical Report.

David, I. P., Molano, W., Fabian, E., Albacea, Z., and Arcilla, R. (2007), "Estimating Food Poverty Directly from Food Consumption Data, 2007," Paper to be presented in the Philippine Statistical Association Anniversary Conference, 19 September.

Derrickson, J., Anderson, J., and Fisher, A. (2000), "Concurrent Validity of the Face Valid Food Security Measure. Institute for Research on Poverty, 2000," *Discussion Paper No.* 1206-00. Available at: http://www.ssc.wisc.edu/irp

Frongillo, E., Rauschenbach, B., Olson, C., Kendall, A., and Colmenares, A. (1996), "Estimating the Prevalence of Hunger and Food Insecurity: The Validity of Questionnaire-Based Measures for the Identification of Households," Final report of the 1994 Small Grants Program by the US Department of Agriculture.

Kendall, A., Olson, C., Frongillo, E. (1995), "Validation of the Radimer/Cornell Measures of Hunger and Food Insecurity," *Journal of Nutrition, 125: 2973-2801.*

Kish, L. (1992), "Weighting for Unequal Pi," *Journal of Official Statistics*, 183-200, Statistics Sweden.

Molano, W., Barba, C., Nueva Espana, Ma. B., and Casio, M. (2002), Household Food Insecurity in the Philippines. Food and Nutrition Research Institute.

Molano, W., Gulles, A., and Tarrayo, Ma. E. (2006), Validation of Food Insecurity Responses vis-a-vis Nutritional Status and Adequacy of Food Intake from the Sixth National Nutritional Survey. Food and Nutrition Research Institute.

Molano, W., Gulles, A., Nueva Espana, M. B., Nones, C., and Tarrayo, M. (2005), Socio-Economic Determinants of Food Insecurity. Food and Nutrition Research Institute.

Molano, W., Nueva Espana, M. B., Tarrayo, E., and Gulles, A. (2004), 2003 Household Insecurity in the Philippines. Food and Nutrition Research Institute.

National Statistical Coordination Board (2007), Available: http://www.nscb.gov.ph/poverty/FAQs/default.asp.

National Statistics Office (2003), 2003 Master Sample Documentation.

Radimer, K., Olson C. and Campbell C. (1990), "Development of Indicators to Assess Hunger," *Journal of Nutrition*, 120: 1544-1548.

Sen, A. (1997), "Hunger in the Contemporary World," *Discussion Paper DEDPS/8*, London School of Economics and Political Science.

Social Weather Stations (2004), Third Quarter 2004 Social Weather Survey Technical Details.

_____ (2007), Available: http://www.sws.org.ph/ind-pov.htm

Tarasuk, V. (2001), Discussion Paper on Household and Individual Food Insecurity. Health Canada.

Webb, P., Coates, J., Frongillo, E. Rogers, B. L., Swindale, A., and Bilinsky, P. (2006), "Measuring Household Food Insecurity: Why It's So Important and Yet So Difficult to Do," *Journal of Nutrition*, 136:1404S-1408S.

Acknowledgements

The author acknowledges Ms. Charissa Castillo and Dr. Isidoro P. David for their valuable comments; and Ms. Wilma Molano, Ms. Carol Duran, Ms. Eleanor Ramos and Ms. Rose Garcia for sharing survey results and technical documentations. This paper was presented at the Philippine Statistical Association Anniversary Conference on 19 September 2007 at the Philippine Social Science Center, Diliman, Quezon City.

Statistical Hypothesis Testing Revisited

Ann Inez N. Gironella

Hello, my fellow teachers. As always, it is a pleasure to share classroom experience with you.

In this article, we will revisit statistical hypothesis testing. This is a procedure used by many experimenters to establish evidence for or against their research hypotheses. In class, when this topic is introduced, students realize the utility of the procedure but many find difficulty in understanding and following the reasoning involved.

So let me begin by saying that the methods of hypothesis testing and of a jury trial are analogous. In statistical hypothesis testing, a null hypothesis is formulated, which is a statement of status quo or no intervention effect (in a jury trial, the null hypothesis is that the defendant is innocent). This is paired with an alternative or research hypothesis, which is a statement of a change in status quo or an intervention effect (defendant is guilty). The end result of a statistical test is action taken on the null hypothesis: either reject it or fail to reject it. The rejection of the null hypothesis in favor of the alternative is based on a *probability value,* more popularly known as *p-value,* of getting a statistic as extreme as what has been calculated if the null hypothesis was true. This statistic is a calculated summary of the data collected for the purpose of the study. Similarly, in a jury trial, declaring the defendant guilty is based on the weight of the evidence against him/her assuming that he/she is innocent. The smaller the p-value, the less tenable is the null hypothesis in light of the data and the stronger is the evidence in favor of the alternative hypothesis. On the other hand, the higher the p-value, the more defensible is the null hypothesis and the weaker is the evidence in support of the alternative hypothesis. Here then are the formal steps to testing a statistical hypothesis.

Step 1: Problem definition

Hypothesis testing entails formulation of the null hypothesis, denoted by H_0, and its accompanying research or alternative hypothesis, denoted by H_a. To introduce this topic to our students, we may ask each of our students to formulate his/her own H_0 vs. H_a. We may start with the following suggestion: take an ordinary coin and test if the coin is fair or not. That is, we ask if the coin is fair or biased. This seemingly simple problem serves as model for experiments with only two outcomes per trial: head or tail in games of chance, yes or no in sample surveys, male or female in gender studies, positive or negative results in clinical trials, defective or non-defective items in production processes. Of interest in problems of this type is the *proportion of successes,* where success is defined as obtaining a favorable outcome.

Step 2: Formulation of H_0 and H_a

We keep in mind that H_0 is generally a statement of status quo (or of equality or no effect or no difference when an intervention is applied) while H_a is a statement of a change in status quo or that an intervention has an effect. H_a is what we would like to establish from evidence presented by the data. For our coin experiment, we expect that our coin is good or fair; this is the status quo. Hence, we formulate H_0: The coin is fair. To translate this hypothesis in statistical terms, we write $H_0 : p = 50\%$, where p is the true but unknown proportion of heads for the coin or probability of heads. What H_a shall we test against H_0? We may argue that with frequent change of hands, the coin may have undergone deterioration and we suspect that it is now biased and take state this as our H_a. On further reflection, however, when we say that the coin is biased, there are three possibilities: (i) it may be biased in favor of heads, in which case $p > 50\%$; (ii) it may be biased in favor of tails so that $p < 50\%$; or (iii) we may be indifferent and simply say that $p \neq 50\%$. We need to identify which one of these three alternatives best expresses our research hypothesis. Our choice of H_a will dictate how the *p-value* will be calculated to arrive at a decision on whether to reject H_0 in favor of H_a. For the sake of discussion let us say we strongly suspect that our coin is biased in favor of heads, that is, $H_a : p > 50\%$.

Step 3: What statistic to use for the test and its distribution under H_0

Since we are testing hypothesis on the proportion, p, of heads turning up for the coin, we know that we can estimate this parameter by the sample proportion, denoted by $\hat{p}$, from a finite number of flips of the coin as

$$\hat{p} = \frac{\text{number of heads turning up}}{\text{number of tosses}} = \frac{x}{n},$$

where x = number of heads turning up in n flips of the coin. Thus, either $\hat{p}$ or x may serve as our test statistic. We also know that for a given number of flips, the sampling distribution of $\hat{p}$ is the same as the distribution of x, which is the binomial distribution with parameters n and p and p = 50% if H_0 is true.

Step 4: Data collection

We then ask the students how we would proceed to collect the data that will allow us to compute the test statistic, $\hat{p}$ or x. It is not unusual to receive many different responses from the students and most likely one would rise to

answer "flip the coin and record the number of heads". Of course, the next question is "how many times do we flip the coin?" At this point, we may accept any or all number of flips suggested so the students may then run the experiment to obtain their data. (See additional note (a) regarding sample size.) Suppose, we have observed and recorded $x = 24$ heads in $n = 33$ flips of the coin. We are now ready to do the test.

Step 5: Computing the *p-value* of the test statistic

The next question now is: could our coin really have been fair in light of what we have observed — 24 heads in 33 flips? Or is this evidence enough to say that our coin is biased? To answer this question, we need to compute the *p-value* of our test statistic. This entails computing the probability of getting 24 or more heads or $\hat{p} \geq \frac{x}{n} = \frac{24}{33} = 72.7\%$, on the assumption that the coin is fair. In other words, we need to compute *p-value* = Prob ($\hat{p} \geq 72.7\%$ when p=50%) = Prob(x $\geq$ 24 when $p = 50\%$). Note that the direction of the inequality follows from H_a. Calculation of this probability yields *p-value* = *0.0088*. (This value may be read off from prepared binomial probability tables for $n = 33, p = 0.50$, and $x \geq 24$ or from statistical software that can compute binomial probabilities). It is interpreted as follows: if the coin is fair, observing at least 24 heads in 33 flips of the coin can happen in only 88 out of 10,000 experiments. What rare occurrence we have observed if the coin is indeed fair! With this as the weight of evidence, we have two choices in regard to deciding what to do with H_0. If we conclude that the coin is fair, we have very little to support H_0 as we have a very slim chance (=*0.0088*) of observing what we have obtained. On the other hand, if we conclude that the coin is biased in favor of heads, and the coin is really fair, the chance that we have made a wrong decision is *0.0088*. Is this chance of an error in decision tolerable?

Step 6: Conclusion

As is customarily practiced, a *p-value* that is smaller than a preset value of 0.01 is deemed very strong evidence against H_0 in favor of H_a while a *p-value* between *0.01* and *0.05* provides sufficient but not strong enough evidence to reject H_0. Since our computed *p-value* is smaller than *0.01*, we conclude that the data provide very strong evidence that the coin is biased in favor of heads. Another way of stating the conclusion is: The observed proportion of heads, *72.7%*, is highly significantly greater than the hypothesized value of *50%*.

Additional notes:

a) Since our conclusions are based only on a sample, we realize that we are liable to committing an error in decision. If the test leads us to reject H_0 when in fact it is true, we have made Type I error. If the test leads us to fail to reject H_0 and H_0 is actually false, we have made Type II error. Unfortunately, we can never tell which of the two errors we may have committed since our decision is based only on a sample. What can be done to reduce the chance of committing these two types of errors, and consequently increase the chance of a correct decision, is to increase the sample size, n. However, increasing the sample size, increases the resources needed to draw conclusions from the study. The problem of determining an appropriate sample size to test H_0 vs. H_a within tolerable levels of Type I and Type II errors for single proportion or one mean is discussed in introductory statistics textbooks.

b) The preset, tolerable, chance of making a Type I error is also known as the level of significance of the test and is denoted by the Greek letter α. The chance of making a Type II error is denoted by the Greek letter β.

c) A test with high probability of rejecting H_0 when it is actually false is said to be a powerful test. It is desirable to employ a test with high power (very close to 100%). Again, in practice, researchers report the power of the test for changes in the status quo most meaningful to their study. In our coin experiment, the power of the test when the coin is biased with $p = 70\%$ for $\alpha = 0.04$ is calculated as: Power = P(rejecting H_0 when $p = 0.70$) = P($x \geq 22$ when $p = 0.70$) = 0.7334. This means that when the coin is really biased favoring heads 70% of the time, the test can detect this bias with probability 0.7334. This is a test with relatively good power.

A final remark: Statistical hyphothesis testing can never establish the truth or falsity of H_0; it can only provide a measure of evidence (*p-value*) against it.

Aims and Scope

The Philippine Statistician (Journal) publishes articles on statistical methods and theory, with more emphasis on applications.

In addition to refereed articles, the Journal will have the following sections that may appear in some of its issues (but not necessarily in all):

Letters to the Editor will provide a forum for the airing of opinions on issues pertinent to the statistical community or offers commentaries on articles that have appeared in the journal.

Notes section will include notices and announcements of upcoming events, conferences, calls for papers, or any concern that can be considered news for the statistical community.

Review section will present reviews on statistics books and software.

Submission of Manuscripts

Submissions should be sent to the Philippine Statistical Association's e-mail (psa.sec@gmail.com) in MS Word format. The manuscript that will be submitted should not have been published elsewhere. Neither should it be under consideration by any other journal.

All manuscripts are refereed and evaluated on content, language and presentation.

A printed page in the Journal will have a maximum amount of space of 4.5" by 8.5".

Organization of Manuscripts

Manuscripts must be organized in the following manner:

- *Title Page*. This should be brief and concise.
- *Author's Footnote*. The author's footnote should present the author(s)'s position, employer, and complete address (including email address).
- *Abstract and Key Words*. An abstract of at most 400 words must be submitted with the manuscript. It precedes the article text. The abstract should summarize results, topics discussed, and main conclusions, but it should not contain any display or complex mathematical notation and no references. If an abstract exceeds the word limit, it will be edited to meet the length restriction.
- *Article Text*. Sections should be numbered; sub- and sub-subsections may be used. Tables, figures, and artwork may be used within the body of the article. A table should be numbered. The original files for the tables, figures, and artwork should also be submitted to facilitate typesetting. Authors must obtain written permission to reproduce or adapt all or part of a figure from a copyrighted source. Mathematical equations cited in the text should be numbered. Numbers should be placed at the rightmost margin of the equation line. Matrices should appear in italics and vectors in bold. All other symbols should appear in italics. The preferred softwares for equations are Mathtype and MS Equations, which are add-ins of MS Word.
- *Acknowledgments*. An acknowledgment section may be included at the end of the article. This section should acknowledge financial assistance in the form of grants or university funding, assistance by individual colleagues, and any other pertinent information. This section will be removed by the editor in the blinded copy of the manuscript.
- *Appendices*. A single appendix is headed, "APPENDIX: FOLLOWED BY A DESCRIPTIVE TITLE". If there are two or more appendices, they should be labeled, "APPENDIX A", "APPENDIX B", and so on.
- *References*. All references included in the References at the end of an article must be cited in the text of the article. For particulars about the style, please download the Philippine Statistician Style Guide at www.philstat.org.ph.

Editorial Style

In addition to content, manuscripts are evaluated on their conciseness and clarity. Thus, the Journal gives premium to well-written and well-structured papers that will be of interest to a wide segment of the readership. Manuscripts and reviews that have been accepted for publication will be copy edited in accordance to accepted rules of correct grammar, usage, spelling, and punctuation. To avoid common problems of style, for guidelines on style, usage, and the preparation of technical manuscripts for publication, the following reference may be consulted:

The Chicago Manual of Style (14th ed.) (1993), Chicago: University of Chicago Press.

Blinding Procedures

The editor will request at least two referees to evaluate a manuscript. The evaluation procedure will be double-blind. Referees will not know the author(s) of the manuscript and the identities of the referees will not also be known to the author(s). In this regard, the editor will make sure that all the references or details that can identify the authors in the manuscript will be removed in the "blinded" version. The title page and/or the front matter of the blinded version of a manuscript should contain no references to any author or to his/her affiliations.

All correspondences to the author regarding the review process should not contain details that can identify the referees.

CONTENTS

The Philippine Statistician
Vol. 57, 2008 © Philippine Statistical Association, Inc.
ISSN 2094-0343