Credit, Insurance, and Impediments to Trade: Micro Underpinnings for Macro Models

MICRO UNDERPINNINGS -- RISK–SHARING --The empirical literature on complete markets and risk sharing began with theory -- Diamond (1967), Scheinkman (l984), and Wilson (l968). Early empirical contributions include Leme (l984), Mace (l991), Cochrane (l991), Altug and Miller (l990), and Altonji, Hayashi, and Kotlikoff (l992), Townsend (l994), and Udry (1994). That literature is expanding rapidly in both developed and developing country contexts -- e.g., Atkeson and Ogaki (India, 1996), Asdrubali (Spain, 1996), Atanasio and Davis (US, 1996), Banks, Blundell, and Brugiavini (Britain, 1999), Deaton (Cote d'Ivoir, 1990), Fafchamps and Lund (Philippines, 1998), Grimard (Cote d'Ivoire l997), Jacoby and Skoufias (India, l997), Jappelli and Pistaferri (Italy, 1998), Okagi and Zhang (Pakistan, 1998), Paulson (Thailand, 1994), Rai, Topa and Amin (Bangladesh, l998), Rashid (Pakistan, l990) and Townsend (Thailand, 1995 a & b).

We begin then with the basic yet strong implications of the full insurance and complete neoclassical model for groups. Suppose we consider individual k in some group v, where this could be a kinship group or local network, some village, or some larger regional aggregate. The appropriate group, if any, will be determined as part of the analysis. Then the basic first-order conditions for individual k from the program delivering full insurance are

$$(1) \quad \lambda^k \; W_c^{\,k}[c_t^{vk}(h_t), l_t^{vk}(h_t), A_t^{vk}] = \mu_c^{\,v}(h_t), \quad \text{all } k$$

$$(2) \quad \lambda^k W_l^{\,k}[c_t^{vk}(h_t), l_t^{vk}(h_t), A_t^{vk}] = \mu_l^{\,v}(h_t), \quad \text{all } k$$

where, in streamlined notation, W_c is the derivative with respect to argument c of the utility function of individual k, in some household j; c_t and l_t are consumption and leisure assignments for individual k; A^k_t is an age-sex index for individual k; λ^k is a weight reflecting the wealth or status of individual k; and μ_c and μ_l are common group v Lagrange multipliers. All choice variables are indexed by the history h_t of economic and demographic shocks. Suppose utility functions are separable in consumption and leisure, with that of consumption taking the form

$$(3) \quad U^k(c_t^{vk}, A_t^{vk}) = -\frac{1}{\sigma_k}\exp\left[-\sigma_k \frac{c_t^{vk}}{A_t^{vk}}\right]$$

Then assuming equal weights λ^k in the household j, identical coefficients of risk aversion σ for all individuals k in all households j, taking time differences to eliminate household fixed effects, and assuming some measurement error u_t^j in consumption, we get the regression equation:

$$(4) \quad c_{t+1}^{vj} - c_t^{vj} = \beta\left(\bar{c}_{t+1}^{v} - \bar{c}_t^{v}\right) + \delta\left(\tilde{A}_{t+1}^{j} - \tilde{A}_t^{j}\right) + \xi\left(X_{t+1}^{j} - X_t^{j}\right) + u_t^j$$

In equation (4), the time difference in consumption of household j in group v is related to the time difference in aggregate group v consumption, a time difference in relative demographic terms of $\tilde{A}_t^j$ picking up the log weighted household j age/sex composition relative to that of the group, and changes in other right-hand side variables X_t^j, for example, the change in household income. The theory implies that the first term, the fixed effect for the group, should be positive, with coefficient $\beta = 1$ picking up the effect of aggregate, group-level risk. No variable in X^j_t (other than aggregate leisure if preferences are non-separable) will enter, not indicators for particular economic shocks (crop disease, water problems, pests, increase in input prices, decline in output prices), not indicators for demographic shocks (episodes of household sickness, changes in days of unemployment, other reasons for not working), nor change in income from all sources (or any source separately -- crops, livestock, enterprise, labor supply). Household specific demographic terms may enter in the $\tilde{A}_t^j$, and also enter in additional terms if there are economies-of-scale in household production. For alternative CRRA preferences the regression equation appears in logs. Regressions in levels, not first differences, will allow estimates of intercepts -- these are closely related to the household's relative wealth or Pareto weight and can be estimated with sufficient data.

It is critical of course that the variables of these equations be well-measured and that spurious effects be avoided. For example, actual average group consumption can be used only with care--as the average of the dependent variable in a cross-section, the coefficient on it would be one, and if the household is left out of the sample average, one in the limit as sample size increases. See Deaton (l990). With sufficiently long time series, these regressions can be run for households individually, allowing, in fact, variations in coefficients of risk aversion, and, as an alternative hypothesis, varying degrees of imperfect insurance; e.g., Townsend (l994) using 6-10 years of annual ICRISAT data. Still, the average of coefficients on group consumption would be mechanically driven to one as sample size increases (see Deaton (1990). Instead, deviations from group average consumption can be taken as the dependent variable, or common time, fixed effects can be used. The latter also corrects for potential bias under the alternative hypotheses there is not perfect risk sharing and there are common components in income. The income coefficient ξ is biased downward in this alternative the larger the correlation in individual incomes. See Ravallion and Chaudhuri (l997). More generally, large standard errors may give the analyst little power to reject anything, and this needs to be checked; for example, see the tables in Townsend (l994). Measurement error in income is a large concern -- with a common effect in actual household incomes (or expected future income) but large measurement errors in observed household incomes, consumption would comove in the data, but not because of insurance, and the coefficient on income would suffer attenuation bias; see Alderman and Paxon (l992). A Griliches/Hausman (1986) test for measurement error in incomes can be employed, utilizing a comparison of first differences in time versus "within" differences from the sample average; e.g., Townsend (l994). Independent checks on the variance, covariance matrices of a crop in various soils, of various crops, and of income from various sources are informative of idiosyncratic and aggregate shocks, e.g., Townsend (1994). Our own direct measures of soil characteristics in the Thai annual data and of soil, rainfall, water flow, and soil stress in the monthly data should help somewhat to sort out measurement error from physical processes. Econometric instruments may also be needed if the measurement error in consumption is also in income, as seems plausible with home produced and consumed items (despite efforts to measure both well in the interviews). The logical instrument for income is income less imputation for home produced goods and other income in kind, as in Deaton (l990), or more radically, income other than from agriculture, as in Ravallion and Chaudhuri (l997). Orthogonal preferences shocks are allowed by these specifications, but one should be careful with shocks such as illness which arguably might lower the marginal utility of consumption.

The questionnaires and preface to the monthly survey contain the details on how consumption and income are actually measured. Suffice it to note as an example that for the monthly instrument the list of commodities for consumption is the one used in the widely accepted annual SES survey. A distinction is made, for example, between milled sticky rice and milled regular rice, and so on, but in our survey we ask about many of these items on a *weekly* basis. Income for the monthly instrument is measured in the individual modules. For example, a given plot in a given crop triggers monthly questions about all inputs and outputs. The annual data is of necessity less accurate. For example, we predict total consumption from a subset of judiciously chosen items with weights from the SES data.

Continuing to assume utility functions are separable in consumption and leisure as above, analogous results apply to leisure. That is, an equation similar to (4) applies for individual labor supply, at least for those individuals never sick or otherwise not hitting binding corner constraints, though it may be necessary to distinguish labor by type (education, disability). Thus changes in the leisure of all laborers of a given type should move together with aggregate (group) leisure and not be influenced by other right-hand side variables, e.g., a declining monthly rice stock position of that household. See also Kochar (1993). More generally, though, separability need not be imposed in analysis of either consumption or labor supply if adequate data are available. See, for example, Altug-Miller (l990), Altonji-Hayashi-Kotlikoff (l996), and Chiappori (1999). Labor supply is reasonably well measured in the monthly data (see further discussion below).

One point to be stressed is that macro aggregates conceal much underlying diversity. Households in the resurvey of May l998 were asked whether the last 12 months were worse in income that the prior year. The answer was in the affirmative for 47% of the households, but there was significant interregional variation. For example 60% and 30% of the surveyed households in Buriram and Srisaket, both in the northeast, reported declining incomes. Similarly, a substantial number of households reported increased incomes – almost one-quarter of the sample overall. Apparently, those with high levels of income from business or farm activities in the base year had higher decreases in income (or lower increases). But, again, Buriram and Srisaket offer a striking contrast, with the histogram of income changes in Srisaket showing an evident right shift. Surprising perhaps, given the reports of unemployment and lost income, households with a large amount of wage income in the base year did reasonably well. Those who did the best were engaged in fish or shrimp farming. In preliminary income-change regressions common village or tambon (sub-county) effects are somewhat significant, but the explanatory power of the regressions plummets when we fail to take into account the income situation of the households in the initial, base year. In other words, there is much diversity over households even within a village, and this should be taken into account in gauging the impact of the crisis (or in policies to provide safety nets or other remedies.)

What were the shocks? Households were also asked to list the three most important causes of income shortfalls. Drought was a big problem for 56% to 77 % of the households with Srisaket the exception at (only) 35%, but the latter changwat in turn suffered more from floods at 24%. This kind of variation is typical of semi-arid climates. Pests were a problem for 32% of Chacherngsao farmers. The macro crisis shows up in complaints about low ouput prices and high input prices at 16% and 21% overall, and in lost wage income in Lopburi at 26%, though 11% overall. Shocks related to illness and associated expenses were important for 5-6% of the population.

Households were also asked how they coped with shortfalls in income (for those claiming to have had a bad income year). Sixty-eight percent of central area households reported reduced consumption expenditures and 34% in the northeast. Preliminary runs of the risk-sharing regressions (4) are consistent -- the coefficient of the impact of idiosyncratic income changes on household consumption change is relatively high and significant in the central area near Bangkok, especially in Chachoengsao, and relatively low and sometimes insignificant in the relatively poor rural northeast. Within a changwat, tambon or village fixed effects can be significant, but the ability to distinguish village from tambon fixed effects is not great. It is as if the average income of villages within a tambon had moved together. Indeed, with the village as the unit of analysis it seems that there is something close to full insurance within a tambon. But that comes apparently from inappropriate aggregation. Split samples are showing that households headed by females, by those over the median age (50), and by those without occupation (retired or disabled) were more likely to have felt the consumption impact of income fluctuations.

MICRO UNDERPINNINGS -- EFFICIENCY IN PRODUCTION --Similarly, when markets or institutions are complete, at least within a group, then production decisions separate from household consumption decisions and household demographics. If we take all farmers in a group v producing rice on similar plots, for example, then we get:

where, in stream-lined notation, the wage is P_{at} at t, the output price is $P_{q,t+1}$ at t+1, μ_c is the Lagrange multiple referred to earlier—hence linking production to consumption aggregates -- $h_t = (\varepsilon_1, \ldots, \varepsilon_t)$ is the history of shocks up to t at the time of decision on labor a_{it} on plot i, and $q_{i,t+1}$ is output at date t+1 on plot i as a function of history h_t, future idiosyncratic shocks $\varepsilon_{i,t+1}$ and aggregate shocks η_{t+1}. If the technology is common across plots, or at least plots of a given type, and if idiosyncratic shocks $\varepsilon_{i,t+1}$

$$\mu_c^v(h_t)P_{at}(\eta_t) = \sum_{\varepsilon_{t+1}} \mu_c^v(h_t, \varepsilon_{t+1})P_{q,t+1}(\eta_{t+1}) \frac{\delta\, q_{i,t+1}\lfloor a_{it}(h_t), \varepsilon_{i,t+1}, \eta_{t+1}\rfloor}{\delta\, a_{it}(h_t)} \qquad (5)$$

across these plots after labor decisions are nonexistent or at most iid, then labor use on the plots of the group should be identical

Benjamin (l992) tests this null of neo-classsical efficiency for rice farmers in Java, using the SUSENAS l980 survey, against the alternative that demographics matter. He postulates that the correct

(6)

shadow wage w* implicit in the household decision is multiplicatively related to the actual market wage—the coefficient is unity (for the null hypothesis) plus a mark up or discount related to household demographics. This delivers, in the context here, equation

$$\log\ a_{it}^{vj} = \left[\alpha \ \ + \beta \log\ \mathrm{P}_{at}\left(\eta_{\mathrm{t}}\right)\right] + \gamma \log\ L_{it}^{j} + \beta \sum \delta_k d_{kt}^{j} + \zeta_t^{j}$$

where a_{it}^{vj} is total man-days on household j's plot i at date t, $L^{\mathbf{j}}_{\mathbf{it}}$ is the amount of land of plot i for household j, and d_t^j is a set of demographic variables for household j, and j is a member of group v. One can then test to see if the number of household members, of children under 15, of prime age males, of elderly males, of prime age females, and of elderly females influence labor use on the plot. (These can be tested as a group.) The specification here should also allow a fixed effect $\alpha^v{}_t$ for group v, capturing group demographics that determine a common shadow wage; with the inclusion of this effect, again, no individual demographics should enter. In the end we have an equation similar to (4) for consumption (and leisure); the group fixed effect can be significant, but the idiosyncratic demographic terms should not be.

Various econometric issues are spelled out in Benjamin's work. First, the functional form may not be correct, but one can move to other specifications, allowing pesticide and fertilizer prices, the wage, area planted, and irrigation status to enter with squared and interactive terms and interacting the wage with demographics. The wage in (6) may be imagined to be measured with error – one remedy is to use the average of measured wages across nearby farmers (village or cluster) as an instrument. The demand for labor in (6) should depend as well on the underlying productivity of the plot – inclusion of soil quality variables should help as controls. Related, the wage in (6) is logically not exogenous to demand but rather determined in a simultaneous equation system. Variables on the supply side such as local population density and presence of a nearby large city can be included in the initial stage of the 2SLS regressions. A Wu-Hausman test statistic indicates the significance of bias in the OLS specification given consistency of the 2SLS specification. Household size may be positively and economically related to labor demand, and the usual simultaneity problem would bias the coefficient on family size in (6) upwards. Alternatively, the family labor component of labor demand could be measured with an error related to family size. More generally, household size may be badly measured and/or correlated with other unobserved and omitted variables, such as land quality, that affect labor demand. Benjamin allows other controls such as age and education of the head in his full OLS regression; uses alternative measurement of family size; allows a distinction between male and female labor; excludes ploughing and harvesting labor as these are related to capital inputs and contractual share-cropping, respectively; and allows areas harvested to be endogenous with demographics.

Related is the work of Feder et al. (1989) on the productivity of wheat farmers in China (see also Feder, et al. 1988). Their particular interest, as here, is whether credit access facilitates neoclassical efficiency. A two-stage switching regression is employed, with stage one determining who (claims) to be constrained and stage two determining output as a function of inputs and possible demographic terms. As in Madalla (1983), the estimation allows correlated errors, that is, allows the likelihood of credit access to be related to farm productivity.

In these micro tests of efficiency in production, we will not focus exclusively on rice in villages, as in the semi-arid northeast. Shrimp farms in Chachoengsao near Bangkok are non-trivial in revenue and frequency. This is a major business leading to a major export. An interesting literature links firms and investment to the informal sector, e.g., Avery and Samolyk (1998), McMillan and Woodruff (1996), Peterson and Rajan (1994, 1997), Schaller (1993), and Zakrajsek (1994). The exhaustive measurement in the Thai survey integrating the household and business survey will allow these and other micro tests and major steps forward in analysis and understanding.

EVALUATION OF INFORMAL AND FORMAL INSTITUTIONS USING RISK METRICS FROM THEORY--There follows a series of explicit tests which make use of these basic regression equations in consumption, leisure, and production in order to analyze the role of informal and formal institutions.

We come for example to the consumption and production effect of being in a formal community organization or, better put, of being in a village with a community organization. Concentrate first on Production Credit Groups (PCG's) and the annual data as a leading example. The most obvious thing to do is to split the annual panel into two strata, those villages with a PCG and those without. Now run regressions such as (4) and their modifications for each sub-sample and compare coefficients for larger fixed effects and smaller idiosyncratic terms. Alternatively, one can pool the data and run one regression with additive and multiplicative terms -- allowing greater sharing of common risk and more diversification of idiosyncratic risk for those in a PCG village. One problem is that we may falsely attribute to the PCG something which is true of the village -- maybe PCG's work well in villages where risk sharing otherwise works well. But at least a correlation would have been established. See Kaboski and Townsend (2001) for more details for more details on how to control for village placement and household selection and estimate the impact of PCG with the 1997 survey.

Institutional affiliation is established in the annual surveys by a series of simple questions. In particular, the household is asked both whether it is now a member and whether it has ever been a member of a PCG. In the preliminary analysis with two years of annual data, we partitioned the sample into those households saying yes to PCG membership both years (thus not identifying PCG villages but rather PCG households). The results are 'perverse' for all four of the changwats, for specifications both in levels and logs. That is, the income coefficient ξ in (4) is higher for PCG members. Evidently there may again be selection effects at work. We are pursuing this. For example, from the institutional questionnaire of the base line we have identified PCG's which have succeeded in expanding membership, savings, and loan size. From the household survey we find that the corresponding villages are ones with small median loans; in which debt/income ratios are low, at least for the relatively poor; and in which the proportion of debt supplied by the informal sector and moneylenders is high. A consistent hypothesis, then, is that these villages still lack adequate credit and insurance facilities. Thus the PCG membership variable would serve more as an index of that underlying 'need' than an indicator of the impact of PCG services themselves**.**

An evaluation of commercial banks follows similarly. It appears that there is a strong association of bank clients with consumption smoothing. The income coefficient ξ is relatively high and significant for non-clients and low and/or insignificant for clients. A notable result for the BAAC is that village, tambon, or changwat effects appear more often in the non-BAAC sample, whereas the BAAC sample (in the log specification) has a common constant in consumption growth. This would be consistent with BAAC members benefiting from a national-level intermediation system, and nonmembers being more affected by local shocks. Townsend and Yaron (2001) take a closer look at the BAAC risk contingency system.

MEASURING ACTUAL ASSET CHANGES — THEIR USE IN TESTS -- Now, as in Townsend (1995b) on northern Thai villages and the work with Lim (1998) using Indian ICRISAT data, we embed these risk-sharing, leisure, and production decisions into the larger household, group, village, or regional problem and begin measurement of how smoothing is actually accomplished or better put, avoiding statements about causality, deficits and their decomposition. Suppressing the group v notation, let the household j budget constraint appear as

$$\begin{aligned} &P_{qt}c_t^j + P_{bt}b_t^j + P_{at}a_t^j - P_{qt}q_t^j - P_{at}\left[T\left(\varepsilon_t^{Hj}\right) - l_t^j\right] \le P_{kt}\left[K_t^{oj} - K_t^{ij}\right] + \left[(1-\delta)S_t^j - S_{t+1}^j\right] + P_{Mt}\left[M_t^j - M_{t+1}^j\right] \\ &+ \left[B_{t+1}^j - \left(1 + r_t^b\right)B_t^j\right]P_{Mt} - \left[L_{t+1}^j - \left(1 + r_t^l\right)L_t^j\right]P_{Mt} + P_{Mt}G_t^j \end{aligned} \tag{7}$$

The left-hand side of (7) is a deficit, the gap between expenditures of all kinds (on consumption c, labor inputs a, and nonlabor inputs b) and income from all sources (revenue from output q and net labor supply). The right-hand side represents a series of mechanisms for smoothing the deficit: (in order) sales of real capital assets K, decumulation of storage S, decumulation of money balances M, increased revenue from borrowing B, decreased outflow of revenue from lending L, or the receipt of gifts and remittances G.

As in the work with Lim and the ICRISAT data, one or several of these mechanisms may track the deficit well. For example let D^j_t denote the deficit at various dates t for household j, and let ΔS^j_t denote the corresponding decrease in crop inventory. We take our measure of tracking of crop inventory to be:

$$\sum_{t=1}^{T}\left(D_t^j - \Delta S_t^j\right)^2 \Big/ \sum_{t=1}^{T}\left(D_t^j\right)^2 \tag{8}$$

where T is the number of observations. This metric is the relative mean square error in tracking, something like 1-R-squared, with zero as a perfect score.

Now, as with Lim, we can use Tukey diagrams to see which type of households, villages, and regions use the various mechanisms. In the ICRISAT data salient patterns were evident, giving us important information about micro underpinnings. The risk-sharing apparent in monthly and annual data, as in the Townsend Econometrica paper, is achieved by the buffer stocks of currency and crop inventory, also "community" institutions providing credit and insurance, but not at all by the purchase and sale of real capital assets and livestock. The buffer stock result complements the related analysis of Chaudhuri and Paxon (1994). We note in particular that currency is used on a year-to-year basis. The use of credit and gifts lends support to the notion that households are not, however, entirely on their own but rather, are linked to one another. The nonuse of livestock and real capital assets contradicts the supposed underpinnings of Rosenzweig and Wolpin (1993). There are also striking patterns by income or wealth, with the poor using currency and the rich using crop inventory, and by village, with the more traditional village using credit less. The pilot Thai project, Townsend (1995b), suggested in contrast that local insurance might deteriorate with the level of development, a common hypothesis in the literature. See Banerjee and Newman (1995), Plateau and Abraham (1991), and Popkin (1979).

The new comprehensive Thai data offers similar possibilities for measurement, but here we have more villages, these villages are in regions of the country which vary in their level of development, and the country itself experienced during data collection both a financial crises and a drought. We bend over backwards in data collection to measure changes in assets, borrowing, and saving, and to distinguish the various informal and formal institutions. Crop inventory is treated in the annual data as a savings account, and in the monthly data we distinguished inventory of milled regular rice from unmilled regular rice, and so on. Weight in kilograms is recorded and put on a roster. Each monthly interview asks for weight again. Differences in weight, which could be from natural processes, trigger questions about acquisitions (e.g., purchases, credit received, gifts), changes by transformation (e.g., milling), and disposition (e.g., feed to livestock, lending). As with the ICRISAT transaction files and the work with Lim, currency use is measured indirectly by distinguishing barter from monetary exchange (in all transactions).

In the annual data, and unlike the monthly data, we have only relatively noisy approximations to both sides of equation (7), the deficit and its decomposition. Still, the essential idea is that actual asset use, as measured, is a variable which may be used in the analysis, e.g., may be associated with the degree to which households approximate the full risk-sharing standard. For example, returning to a preliminary evaluation of the potential benefits of formal and informal financial institutions, we note that the BAAC, a dominant lender in the 1997 cross-section with 34% of all loans outstanding and 29% by total value, has been relatively inactive in the crisis, with 10% of households showing a decrease in debt and l5% an increase. Then, partitioning the sample into three strata, those with decreases in debt, those with no change (and possibly not a member), and those with increases, we are rerunning the risk-sharing regression (4). Early results show that households with no change in BAAC debt have been more vulnerable to risk in the sense that the coefficient ξ on changes in household income is higher, but those with reduced BAAC debt sometimes suffered more. Similarly, those with no change in commercial bank savings did worse in shedding risk on average, but, in contrast, as in a buffer stock model, those who reduced commercial bank savings often did better.

Of great interest is the role (impact or association) of the informal sector. Informal debt levels, a nontrivial 40% of all loans in the base line by value, decreased in the resurvey for 14% of the population and increased for a relatively large 34%. Still, preliminary results in the risk-sharing regressions show in more than half of the cases that income coefficients are particularly high for those with increases in informal debt. One interpretation is that households under stress and with insufficient savings fail to get sufficient flexibility in loans from the formal sector and find actual increases in informal debt inadequate ex post to cover income fluctuations. If true, this would be consistent with earlier work on tiered credit markets (e.g., Kochar (1997b), and Bell, Srinivasan and Udry (1992)).

IMPEDIMENTS TO TRADE -- PERMANENT INCOME, BUFFER STOCK, AND PRIVATE INFORMATION MODELS--Noting that the full risk-sharing benchmark is often rejected, we turn next to alternative and more limited standards for consumption smoothing. In the permanent income model, insurance and gifts are exogenously set to zero. Alternatively, in the buffer stock model, there is savings but borrowing is at best bounded. The empirical implications of the permanent income and buffer stocks models are, of course, well known, e.g., Banks, Blundell and Tanner (l997), Carol and Samwick (1998), Deaton (l989, 1991), Flavin (l983), Hall (1978), Hansen and Singleton (l983), Kimball (1990), Lim (l993), Paxon (l992), Wolpin (l982), and Zeldes (l989), just to name a few. In a mechanism-design world a lender controls the borrower's assets and uses consumption rewards for incentive purposes. The implication of this for borrowing/lending and consumption is discussed in Allen (1985), Atkeson and Lucas (1992), Fudenberg, Holmstrom, and Milgrom (1990), Green and Oh (1991), Ligon (1998), Phelan and Townsend (1991), Rogerson (1985), Spear and Srivastava (1987), and Townsend (1982).

To be more specific about tests, imagine a permanent income world with perfect internal or external credit markets and common real borrowing and lending rates r_t at date t to date t+1 (and either no other assets or equality in rates of return). Then we get

$$(9) \qquad U_c^j(c_t^j) \geq E\beta(1+r_t)U_c^j(c_{t+1}^j)\,,$$

with equality for the permanent income world and with inequality when buffer stocks hit zero and the household would like to borrow but cannot. Zeldes (1989) writes down a generalized, empirical version of (9). His first-order condition is of the form

$$(10) \qquad U_c^j\left(c_t^{vj}, \theta_{jt}\right) = E_t \frac{U_c^j\left(c_{t+1}^{vj}, \theta_{j,t+1}\right)\left(1 + r_{jt}\right)}{1+\delta}$$

Here, in stream-lined notation, c_t refers to consumption of household j in group v at date t, θ_j is a demographic household-j- specific shock, δ is a discount rate, r is a realized interest rate (possibly depending on j) and E is an expectation operator relative to all variables known at date t by household j. This leads to an equation like (10) but without expectation and with the inclusion of an error term uncorrelated with information known by the household at time t. The latter may contain an aggregate component as well as a household-specific, idiosyncratic component. Also the household shock θ_j consists of a fixed family effect, an aggregate time group effect, and an orthogonal component as well as systematic change due to age and a measure of family size, AFN. With CRRA preferences to the coefficient (1-α) and the shock entering exponentially, this delivers regression

$$(11) \qquad \ln\left(c_{t+1}^{vj}/c_t^{vj}\right) = k^1 + k_t^{2v} + k_j^3 + \frac{1}{\alpha}\left[\ln\left(1+r_{jt}\right) + b_1 age_{jt} + b_2 \ln\left(AFN_{j,t+1}/AFN_{jt}\right)\right] + v_{j,t+1}$$

On the assumption that income and other variables known at date t should be unrelated to the error terms in this equation, one can then run the regression equation (11) with the inclusion of date t income, or other variables. If the permanent income model is a good approximation, the income term should not be significant. Alternatively, following Zeldes, partition the data, having one group for whom observed liquid assets stay positive (or for whom the wealth to income ratios stay high), and having a second group with zero or low values for these variables. In this case, the income variable should be significant for the second group but not the first.

Under specified assumptions, Ligon (1998) derives an associated Euler condition for a private information moral hazard model:

$$(12) \qquad 1/U_c^j\left(c_t^j\right) = E\left[1/U_c^j\left(c_{t+1}^j\right)\right]$$

With β=1/(1+r), this is precisely the inverse of the usual Euler equation (9) at equality, and thus CRRA preferences allows him to nest the two models—the sign of the estimated preference parameter puts the household in one model or the other. That is, using GMM techniques from Hansen (1982), Ligon can identify in the ICRISAT data households which seem to pass private information tests, and ipso facto fail a permanent income test. Such information-constrained households would be savings-constrained, i.e., would like to carry some buffer but are not permitted to do so. (This is why we ask borrowers if their lender is keeping track of their savings.) If we imagine that currency is hard to control, then households using currency as a buffer should not pass (12) at conventional risk aversion values.

Indeed, to return to the Lim/Townsend criterion for tracking the deficit, the relatively poor in the ICRISAT data tended to be high users of currency, and these and other such users of currency were more likely to fail private information if not other model tests. The relatively rich tended to be high users of crop inventory, and these and other such users of crop inventory were more likely to pass permanent income tests. With the monthly Thai data, we will be able to perform these tests of one model against another – and find possible associations with asset use.

Particularly troublesome are cases where nonuse of a lender or asset is associated with failure of a model test. More fully specified general equilibrium models are needed to make the choice of a lender or market and dynamic asset behavior endogenous. This is the thrust of the equilibrium literature with transactions costs and non-financial income, making endogenous the codetermination of asset-market participation and asset prices, e.g., Constantinides and Duffie (1996), Guiso-Japelli-Terlezzese (1996), Heaton-Lucas (1997a & b), Luttmer (1997), and Vissing-Jorgensen (1999). Below, in the section on macroeconomics and dynamics, we postulate and use one such model, namely Greenwood and Jovanovic (1990) and the extended model of Townsend and Ueda (2001).

ENTREPRENEURSHIP AND OCCUPATION CHOICE--We turn next to the literature emphasizing occupation choice, entrepreneurship, production, and the role of credit. Specifically, the incomplete market model of Lloyd-Ellis and Bernhardt (LEB, for short) (1999) postulates fixed costs for establishment of a firm or enterprise outside the subsistence sector, though these costs vary with the talent (education, human capital) of the household. An alternative to both subsistence and entrepreneurship is to work for wages for the firms themselves. The LEB model essentially shuts down the market for savings and credit exogenously, though one could contemplate what would happen if a sub-sector were given competitive intermediaries (see below). Related are the limited-credit models of Evans and Jovanovic (EJ, for short) (l989) and Holtz–Eakin, et al. (l994a & b). In these, agents can borrow, adding to existing wealth the capital for investment in a production function. In that function, physical and human capital (talent) are complements. Yet agents can potentially renege on debts. If caught, they suffer a penalty. This essentially limits credit in the model as in Banerjee and Newman (1993), and Ghatak, Morelli, and Sjostrom (1998) to a positive multiple of wealth. Related also are the moral hazard models of Aghion and Boulton (l994) and Lehnert (l998), (ABL) for short. In these models, borrowing again increments wealth in financing an observed capital input. That input along with unobserved effort delivers output, but with a positive probability of bankruptcy. The talented have less disutility in effort and so for a given utility level go bankrupt less often. The moral hazard problem in these models can limit credit for some or eliminate it for others entirely. Typically one sees poor households who at best can save their limited wealth, intermediate households who borrow to finance production though at an inefficient or constrained level relative to full information, and rich households who self-finance entirely and put their residual wealth into savings. ABL is representative of a larger class of mechanism design models, e.g. Piketty (1994).

Anna Paulson and I begin by retaining the original structure of the LEB, EJ, and ABL models and by asking how well these models do in general and in competition with one another in the relatively large l997 cross-sectional retrospective survey. Among the 2880 households of the sample, 606 are running businesses -- 28% of the sample in the central region and 13% in the northeast. We go back five years before the survey and look at our retrospective estimates of household wealth at that time, and whether the household was a client or member of a financial institution at that time, and then record whether the household went into businesses between that time and the May l997 survey. And though this is not our favorite variable, we also know whether or not a business is contemporaneously constrained, that is, reporting that it is constrained in profits by limited credit.

All three models can be summarized in the common space of wealth and talent. We thus stratified prior wealth by quartiles. Similarly, we took education as the best available measured proxy of business talent. As the majority of current household heads had the mandatory four-year education of their cohort, we considered that category and two others, one above four years and one below. (We also consider education of the most-educated member and vocational training to be variables.) With these stratifications, we find facts consistent with all three of the models, facts consistent with one or two but not all three of the models, and facts possibly inconsistent with all three models.

All of the models predict, holding business skill fixed, that as wealth increases the percentage of business owners increases. This prediction is strikingly confirmed in the Thai data. For example, for the highest education group, the percentage of business owners increases from 24% to 50% in the central region and from 21 % to 34% in the northeast, as one goes from the lowest wealth quartile to the highest. Similarly, all the models imply that for at least some wealth levels, the percentage of businesses will increase when business talent (education) increases. This also is true of the Thai data.

There are differences across the models. When wealth is fixed and skill increases, the LEB model predicts a decrease in constrained entrepreneurs, and the EJ model, with the notion that human and physical capital are complements, predicts an increase. Comparing the number of constrained business owners in the lowest education category to the highest education category, for a given wealth quartile, we find the percentage of constrained business owners increases in three of the four quartiles in the northeast and two of the four quartiles in the central region. More formal probits for entering business suggest that entrepreneurial talent and wealth have differing roles in the distinct regions of the country.

Age of the head, age squared, number of adult males, and number of females are added as controls in the probits; as noted in the section on efficiency in production, some of these are also the demographic variables that Benjamin (1992) and Feder et al. (1989) predict should not be included if the household business were not constrained. We also include as statistical controls in these probits access to informal and formal institutions, as the LEB model itself has no credit. In another set of probits, we also add these institutional variables interacted with wealth. We then find that the effect of wealth on the incidence of entrepreneurship varies by institution.

In Paulson and Townsend, we estimate the LEB model via maximum likelihood. The point is that once one specifies the parameters of the production function, the subsistence level, the cost of living, and the skewness of talent in the population, the model delivers a likelihood that a household, given observed wealth, would enter self-employment or remain as a worker or farmer. Thus we are choosing parameters to maximize the likelihood that our sample of households would have made the choices we observed in the data. Essentially, all the intra-temporal parameters are identified. Preferences for bequests require information on savings rates. Similarly, we estimate the EJ model. EJ assumes talent and individual shocks are log normal in the population, allows a reduced form correlation between wealth and talent, allows wages and profits to depend on a vector of observed attributes, and imposes an empirically observed interest rate. The estimated parameters include the marginal productivity of capital, the maximum credit/wealth ratio, and the distribution of entrepreneurial talent relative to idiosyncratic variation in earnings. Finally we use a linear program approach to determine the likelihood of the Aghion-Boulton moral hazard model and estimate risk aversion among other parameters.

The larger goal here is to compare formally across models, e.g., LEB versus EJ, but fortunately one can make headway using the maximized log likelihoods of each model. We adopt the procedure of Vuong (1989). We find the ABL model fits best overall, but the EJ model dominates among low-wealth households and dominates in the northeast. Other non-parametric tests support the same conclusion. For example, credit is increasing in wealth among low-wealth households in the northeast as the EJ model predicts. The opposite prediction supports the ABL model among high-wealth households and those in the central region. An evaluation of the credit markets in these various regions and knowledge of the lending procedures of particular institutions such as the BAAC allow us to understand better the context of these results.

MACRO DYNAMICS—Thailand experienced a high growth period, 8.38% GDP growth from l986-l995. This came with increasing (and then decreasing) inequality and substantial financial deepening. In terms of measured financial (macro) aggregates, the ratio of M2 to GDP surpassed that of the US by l992, and in the SES (micro) household surveys, households gaining access to the financial sector increased from 6% to 26% from l976-l996 (though the SES data are noisy on that dimension). In short, inequality movements with financial deepening in Thailand were nontrivial dynamic phenomena.

The LEB model above is in fact a dynamic model that was designed by its authors to generate growth with increasing and then decreasing inequality. Growth is driven initially by capital accumulation, but a large number of subsistence workers keeps wages low and profits and income in the hands of entrepreneurs. Eventually wages rise and inequality is reduced. Further, the model can be modified to include exogenously a nontrivial financial sector in which there is perfect intermediation, with a savings/credit market clearing at a market determined interest rate, and with entry into entrepreneurship determined in that sector exclusively by talent, not wealth. With an exogenous rate of increase in the intermediated sector, financial deepening plays a role in growth.

Gine and Townsend (2001) numerically compute and simulate the dynamic LEB model using the parameter estimates of the maximum likelihood estimation. We are thus using micro data in estimation and looking at the macro outcomes of a model with exactly the micro underpinning as were assumed in the estimation. We are using as an initial condition an estimate of the distribution of wealth, which Jeong (1999) has backed out of the SES l976 data. Otherwise, virtually all the parameters of the LEB model are pinned down from estimation with the project data or SES data, with the exception of the propensity to save and an (exogenous) growth rate for the subsistence sector. These last two can be pinned down by matching the dynamic model economy to savings and growth rates from the national income accounts. More generally, we are following the two steps of the calibration/verification procedure of Christiano and Eichenbaum (l992), CE for short. They are in turn motivated by the real business cycle literature, e.g., Kydland and Prescott (1982). But as emphasized by Hansen and Heckman (l996), HH for short, the CE procedure is actually a well-established econometric procedure: estimation and testing (estimation from some variables or limited aspects of those variables and testing with the other variables or important residual aspects). For example CE pin down (most of) their parameters by some first moment conditions, and then use covariances, the covariation of hours worked and average productivity, to judge the distance of the model from US data. Multiple models are considered, e.g., with and without government and with and without indivisibilities in labor supply. Here we will pin down via formal maximum likelihood estimation, using our own project data and the SES data, (most of) the underlying parameters. We then use an additional criterion to judge the ability of the model at those estimated parameters to explain the actual evolution of the level and distribution of income. In related work, Jeong and Townsend derive kernel density estimates of the income distribution of Thailand from the SES data and from the LEB model (at given points in time), and then use the integrated square error criterion to compare the two. We know from Li (1996) that a test statistic is normally distributed in the limit and so the p-value that the two distributions are the same gives us a reasonable metric. Or we can use the Kramer-von-Mises or Kolmogorov-Smirnov (KS) test – the latter test statistic has a well known distribution, tests are automated in STATA, and the framework can handle spikes in the distribution functions, as in the LEB model for workers receiving a common wage. We can also compare multiple models – for example, the original LEB model versus the one with our exogenously expanding credit sector, or the LEB model against one with endogenous financial deepening (Greenwood and Jovanovic 1990).

As Hansen and Heckman note, the Christiano-Eichenbaum procedure does take into account the precision of estimates of the underlying parameters, and we too need to consider standard error bands around parameter estimates of the maximum likelihood procedure, but CE does not take into account error due to model mis-specification, acting unrealistically as if all error were sampling error. A given model might assume complete insurance, for example, whereas in reality that at best may be an approximation. With even less assurance that estimated parameters represent reality, HH recommend simulations at alternative parameter values, that is, sensitivity analysis. Related, our most recent simulations of the LEB model at the estimated parameter values can match savings rates and seem somewhat consistent with labor shares, but do less well with period by period variation in growth rates. But this is informative. It is important to know on what dimensions a model is failing, not just where it succeeds, one of the few premises held in common by a diverse set of macro practitioners. Or perhaps the model will succeed at alternative parameter values. Watson (l993) creates a new criterion function, the variance of errors which reconcile the variables of a model with those in the data, and then looks at alternative weights on various of the variables and alternative filters or segments of the spectra. In spirit, we can do the same, varying parameters and concentrating on growth rates, for example, trying to explain the sharp upturn in growth in l986. When we adjust growth to remove apparent technological progress, TFP, and use more information on financial deepening, mainly the expansion rate, we match growth reasonably well. On the other hand, returning again to the micro data, but in this case the SES national survey, we will explore the ability of the LEB model to track not only the distribution income but also the contributions to growth and inequality associated with occupation shifts and credit access. As detailed in the thesis of Jeong (l999), occupation shifts, educational attainment, and credit access account from l976-l996 for 39% of average income growth and 53% of changing inequality. But while occupation shifts account for decreased inequality, credit increased it. Jeong and Townsend (2001) pursue this decomposition further using the structure of the LEB model.

The model of Greenwood and Jovanovic (1990) (GJ) is a general equilibrium dynamic model with increasing inequality and endogenous financial deepening which builds directly on the premise that risk-sharing is good (perfect) for those with access to the formal financial system and poor (as if in autarky) for those without access. Entry into the financial system is a function of wealth. This makes the selection of institutions endogenous jointly with patterns in the income/consumption data. This is due to supposed transactions costs, that is, nontrivial fixed and marginal costs of entering and using the (complete markets) financial system. In this GJ model, those in the system take advantage of both improved information and full insurance against idiosyncratic shocks to select production activities. Those outside the system may decide to diversify into low-yield but low-variance activities, as Morduch (1993) suggested for ICRISAT data. We have noted above the indications that these supposed micro underpinnings are validated in an approximate sense in the Thai project data - clients of the BAAC and commercial banks do seem to do better at risk-sharing than others. More generally, the GJ model is generalized in Townsend and Ueda (2001) and its equilibrium computed. We then estimate its parameters using the SES and project data and the stationary, recurrent features of the model. For example, the technology parameters can be estimated with the data on income and assets and preference parameters estimated with data on savings and the share of assets in various occupations. Entry into the financial system is also a stationary stochastic function of past wealth. At estimated parameter values, we simulate the model and deliver growth, inequality, and financial deepening not unlike those observed in Thailand.