A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative

Joseph P. Kaboski and Robert M. Townsend[*,†]

September 6, 2008

Abstract

This paper uses a structural model to understand, predict, and evaluate the impact of an exogenous microcredit intervention program, the Thai Million Baht Village Fund program. We model household decisions in the face of borrowing constraints, income uncertainty, and high-yield indivisible investment opportunities. After estimation of parameters using pre-program data, we evaluate the model's ability to predict and interpret the impact of the village fund intervention. Simulated predictions from the model mirror actual data in reproducing a greater increase in consumption than credit, which is interpreted as evidence of credit contraints. A cost-benefit analysis using the model indicates that the program costs twenty-five percent more than a transfer program providing an equivalent benefit.

[*]This research is funded by NICHD grant R03 HD04776801. We would like to thank Sombat Sakuntasathien for collaboration and making the data acquisition possible. We would like to thank Aleena Adam, Francisco Buera, Xavier Gine, Donghoon Lee, Audrey Light, Masao Ogaki, Mark Rosenzweig, and Bruce Weinberg for helpful comments. We have also benefited from comments received on presentations of versions of this work at IUPUI, NIH, Ohio State, UW Milwaukee, NYU, Yale and the following conferences: NEUDC 2006, 2006 Econometric Society, BREAD 2008, and World Bank Microeconomics of Growth 2008, UC-UTCC. Bin Yu, Taehyun Ahn, and Jungick Lee provided excellent research assistance on this project.

[†]Email: Kaboski <kaboski.1@osu.edu>, Townsend <rtownsen@uchicago.edu>

1 Introduction

This paper uses a structural model to understand, predict, and evaluate the impact of an exogenous microcredit intervention program, the Thai Million Baht Village Fund program. Understanding and evaluating microfinance interventions, especially such a large scale government program, is a matter of great importance. Proponents of microfinance argue that the unique policies of microfinance enable institutions to bring credit and savings services to underdeveloped areas and to people with otherwise insufficient or no access to formal financial systems. The hope and claim is that the provision of saving and credit is both effective in fighting poverty and more financially viable than other means, while detractors point to high default rates, reliance on (implicit and explicit) subsidies, and the lack of hard evidence of their impacts on households. The few efforts to evaluate the impacts of microfinance institutions using astructural methods and plausibly exogenous data have produced mixed and even contradictory results.[1] To our knowledge, this is the first structural attempt to model and evaluate the impact of microfinance. Three key advantages of the structural approach are the potential for quantitative interpretation of the data, counterfactual policy/out of sample prediction, and well-defined normative program evaluation.

The Thai Million Baht Village fund program is one of the largest scale government microfinance initiatives of its kind.[2] Started in 2001, the program involved the transfer of one million baht to each of the nearly 80,000 villages in Thailand to start village banks that are run by a committee of villagers and lend to village members. The transfers themselves

[1] Pitt and Khandker (1998), Pitt et al (2003), Morduch (1998), Coleman (1999), Gertler, Levine and Moretti (2003), and Karlan and Zinman (2006) are examples. Kaboski and Townsend (2003) estimates positive impacts of microfinance in Thailand using non-experimental data.

[2] The Thai program involves approximately \$1.8 billion in initial funds. This injection of credit into the rural sector is much smaller than Brazilian experience in the 1970s, which saw a growth in credit from about \$2 billion in 1970 to \$20.5 billion in 1979. However, in terms of a government program implemented through village institutions and using micro-lending techniques, the only comparable government program in terms of scale would be Indonesia's KUPEDES village bank program, which was started in 1984 at a cost of \$20 million and supplemented by an additional \$107 million in 1987. (World Bank, 1996)

sum to about 1.5 percent of Thai GDP, and substantially increased available credit. We study a panel of 960 households from sixty four rural Thai villages in the Townsend Thai Survey. In these villages, funds were founded between the 2001 and 2002 survey years, and village fund loans amounted to eighty percent of new short-term loans and one third of total short-term credit in the 2002 data. If we count village funds as part of the formal sector, participation in the formal credit sector jumps from 60 to 80 percent.

We view this injection of credit, an initiative of (then newly-elected, now) former Prime Minister Thaksin Shinawatra, as a quasi-experiment that produced exogenous variation over time and across villages. The program was unanticipated and rapidly introduced. More importantly, the total amount of funding given to each village was the same (one million baht) regardless of the number of households in the village. Although village size shows considerable variation within the rural regions we study, villages are administrative geopolitical units and are often subdivided or joined for administrative or political purposes. Indeed, using GIS maps, we have verified that village size patterns are not related to underlying geographic features and vary in bi-annual data. Hence, there are a priori grounds for believing that this variation and the magnitude of the per capita intervention is exogenous with respect to the relevant variables.

Our companion paper, Kaboski and Townsend (2008), examines the exogeneity and impacts of the program using an astructural regression approach. We indeed show that village size is not significantly related to pre-existing differences (in levels or trends) in credit market or relevant outcome variables. After the program, however, many outcome variables are strongly related to village size, and many of these are puzzling without an explicit theory of credit constrained behavior. In particular, households increased their borrowing and their consumption roughly one for one with each dollar put into the funds. A perfect credit model, such as a permanent income model, would have trouble explaining the large increase in borrowing, since reported interest rates on borrowing did not fall as a result of the program. Similarly, even if households treated loans as a shock to income rather than a loan, they would only consume the interest of the shock in the initial period (roughly six percent). Moreover, although the injections eventually led to higher default rates, defaults remained relatively low. Finally, although household investment is an important aspect of household behavior, impacts of the program on aggregate investment were difficult to discern, though some increase in the frequency of investment was observed. This is a priori puzzling if credit constraints are deemed to play an important role.

The structural model we develop in this paper here sheds light on many of these findings. Given the prevalence of income shocks that are not fully insured in these villages (see Chiappori et al., 2008), we start with a standard precautionary savings model (e.g., Aiyagari, 1994, Carroll, 1997, Deaton, 1991). We then add important features designed to capture other key aspects of the economic environment and household behavior in the pre-program data. In particular, short-term borrowing exists but is limited, and so we allow borrowing but only up to limits. Similarly, default exists in equilibrium, as does renegotiation of payment terms, and so our model incorporates default. Finally, investment is relatively infrequent in the data, but is sizable when it occurs. To capture this lumpiness, we allow households to make investments in indivisible, illiquid, high yield projects whose size follows an unobserved stochastic process.[3] Finally, income growth is high but variable, averaging 8 percent but varying greatly over households, even after controlling for life cycle trends. Allowing for growth requires writing a model that is homogeneous in permanent income, so that a suitably normalized version attains a steady state solution, giving us value functions and time-invariant (normalized) policy functions. These features are not only central features of the data, but central to the evaluation of microfinance.

Our approach is indeed to estimate the model in an attempt to closely mimic relevant features of the pre-program data by matching income growth volatility, default rates, saving/borrowing rates, and consumption and investment behavior of people with different levels of observed income and liquid assets. We estimate 10 parameters using a Method of Simulated Moments (MSM) across 15 moment conditions. The model broadly reproduces many important features of the data, including closely matching consumption, investment

[3] An important literature in development has examined the interaction between financial constraints and indivisible investments. See, for example, Banerjee and Newman (1993), Galor and Zeira (1993), Gine and Townsend (2001), Lloyd-Ellis and Bernhardt (2001), and Owen and Weil (1997),

levels and investment and default probabilities. Nonetheless, two features of the model are less successful, and the overidentifying restrictions of the model are rejected. The income process has trouble replicating the data, which are affected by the Thai financial crisis in the middle of our pre-intervention data, and the borrowing and lending rates differ in the data but are assumed equal int he model. Using the model to match year-to-year fluctuations is also difficult.

For our purposes, however, a more relevant test of the usefulness of the model is its ability to predict responses to an increase in available credit, namely the village fund intervention. Methodologically, we model the microfinance intervention as an introduction of a borrowing/lending technology that relaxes household borrowing limits. These limits are relaxed differentially across villages in order to induce an additional one million baht of credit in each village; hence, small villages get larger reductions of their borrowing constraint.

We then simulate the model with the stochastic income process to create 500 artificial datasets of the same size as the actual Thai panel to examine whether the model can reproduce the above impact estimates. The model does remarkably well. In particular, it predicts an average response in consumption that is close to the dollar-to-dollar response in the data. Similarly, the model reproduces the fact that effects on average investment and investment probability are difficult to measure in the data.

In the simulated data, however, these aggregate effects mask considerable heterogeneity across households, much of which is unobservable. Increases in consumption come from roughly two groups. First, there are hand-to-mouth consumers, who are constrained in their consumption either because they have low current liquidity (income plus savings) or are using current (pre-program) liquidity to finance lumpy investments. These constrained households use additional availability of credit to finance current consumptions. Second, households who are not constrained may increase their consumption without even borrowing. They simply reduce their bufferstock savings, which is less needed in the future given the increased availability of credit. Third, for some households, increased credit induces them to invest in their high yield projects. Some of these households may actually reduce their consumption, however, as they supplement credit with reduced consumption in order to finance sizable indivisible projects. (Again, evidence for such behavior in the pre-intervention data is an important motivation for modeling investment indivisibility.) Finally, for households who would have defaulted without the program, available credit may simply be used to repay existing loans and so have little effect on consumption or investment. Perhaps most surprising is that these different types of households may all appear *ex ante* identical in terms of their observables.

The model not only highlights this underlying heterogeneity, but also shows the quantitative importance of these behaviors. Namely, the large increase in consumption indicates the relative importance of the first two types of households, both of whom increase their consumption. Also, the estimated structural parameters capture the relatively low investment rates, and large skew in investment sizes. Hence, overall investment relationships are driven by relatively few, large investments, and so very large samples are needed to measure effects on average investment. Second, given the lumpiness of projects, small amounts of credit are relatively unlikely to change investment decisions, particularly decisions on the large projects that drive aggregate investment.

We use the model and data for several other alternative analyses. First, we re-estimate the model and borrowing constraint parameters using the pooled pre- and post-intervention data, and verify that these estimates are quite similar to our baseline estimates and calibrated borrowing constraints. Hence, the model we use to do predictions is quite similar to a "best fit" model in the overall data. Second, we use the model to simulate long run predictions and show that measured impacts from reduced form regressions fall substantially with the number of years of post-treatment data. Third, we model a counterfactual policy in which the microfinance funds only lend to households that invest in the period when they borrow. Such a policy has larger effects on investment than the implemented policy but still not as large as the effect on consumption.

Finally, our normative evaluation compares the costs of the Million Baht program to the costs of a direct transfer program that is equivalent in the sense of providing the same utility benefit. The heterogeneity of households plays an important role and indeed the welfare benefits of the program vary substantially across households and villages. Essentially, there are two major differences between the microfinance program and a well-directed transfer program. First, the microfinance program is potentially *less* beneficial because households face the interest costs of credit. In order to access liquidity, households borrow more, and while they can always carry forward more debt into the future, they are left with larger interest payments. This is particularly harmful for defaulting households. On the other hand, the microfinance program is potentially *more* beneficial because it can also provide more liquidity to those who potentially have the highest marginal valuation of liquidity by lowering the borrowing constraint. Quantitatively, given the high level of default in the data and the high interest rate, the first effect dominates and the transfer program is more cost-effective, costing on average only 80 percent what the microfinance program cost. Nonetheless, the program is relatively more cost-effective for non-defaulting households with urgent liquidity needs for consumption and investment.

The paper contributes to several literatures. First, we add a structural modeling approach to a small literature that uses theory to test the importance of credit constraints in developing countries (e.g., Banerjee and Duflo, 2002). Second, we contribute to an active literature on consumption and liquidity constraints, and the bufferstock model, in particular. Studies with U.S. data have also found a high sensitivity of consumption to current liquidity (e.g., Zeldes, 1989, Aaronson, Agarwal, and French, 2008), but we believe ours is the first to study this response with quasi epxerimental data in a developing country. Third, methodologically, we build on an existing literature that has used out of sample prediction, and experiments in particular, to evaluate structural models (e.g., Todd and Wolpin, 2003). Finally, we contribute to the literature on measuring and interpreting treatment effects (e.g., Heckman, Urzua, and Vytlacil, 2004), which has emphasized unobserved heterogeneity, non-linearity and time-varying impacts. We develop an explicit behavioral model where all three play a role.

The remainder of the paper is organized as follows. The next section discusses the underlying economic environment, the Million Baht village fund intervention, and reviews the facts from reduced form impact regressions that motivate the model. The model is presented in Section 3, while Section 4 summarizes the computational methods used to solve the model, and the resulting value and policy functions. Section 5 discusses the data and presents the GMM estimation procedure and resulting estimates. Section 6 simulates the Million Baht intervention, performs policy counterfactuals, and presents the welfare analysis. Section 7 concludes.

2 Thai Million Baht Credit Experiment

The exogenous intervention that we consider is the founding of village-level microcredit institutions by the Thai government, the Million Baht Fund program. Former Thai Prime Minister Thaksin Shinawatra implemented the program in Thailand in 2001, shortly after winning election. One million baht (about $24,000) was distributed to each of the 77,000 villages in Thailand to found self-sustaining village microfinance banks. Every village, whether poor or wealthy, urban[4] or rural was eligible to receive the funds. The size of the transfers alone, about $1.8 billion, amounts to about 1.5 percent of GDP in 2001.

The design and organization of the funds were intended to allow all existing villagers equal access to loans through competitive application and loan evaluation handled at the village level. Funds were disbursed to and held at the Thai Bank of Agriculture and Agricultural Cooperatives, and funds could only be withdrawn with a withdrawal slip from the village fund committee. Village fund committees were large (consisting of 24 members) and representative (e.g., half women, no more than one member per household) with short, two year terms. Residence in the village was the only official eligibility requirement for membership, and so although migrating villagers or newcomers would likely not receive loans, there was no official targeting of any sub-population within villages. Loans were uncollateralized, though most funds required guarantors. There were no firm rules regarding the use of funds, but reasons for borrowing, ability to repay, and the need for funds were

[4]The village (moo ban) is an official political unit in Thailand, the smallest such unit, and is under the province (changwat), district (amphoe), and sub-district (tambon) levels. Thus, "villages" can be thought of as just small communities of households that exist in both urban and rural areas.

the three most common loan criteria used. Indeed many households were openly granted loans for consumption. The funds make short-term loans – the vast majority of lending is annual – with a typical nominal interest rate of six percent (about 4 percent real).[5]

2.1 Quasi-Experimental Design of the Program

As described in the introduction, the program design was beneficial in two ways. First, because it arose from a quick election, after the Thai parliament was dissolved in November, 2000, and was rapidly implemented in 2001. None of the funds had been founded by our 2001 (May) survey date, but by our 2002 survey, each of our 64 village had received and lent funds, lending 950,000 baht on average. Households would not have anticipated the program in earlier years. We therefore model the program as a surprise. Second, the same amount was given to each village, regardless of the size, so smaller villages received more funding per household. Regressions below report a highly significant relationship between household's credit from a village fund and inverse village size in 2002 after the program.

There are strong *a priori* reasons for expecting variation in inverse village size to be fairly exogenous with respect to important variables of interest. First, villages are geopolitical units, and villages are divided or merged based on fairly arbitrary redistricting. Second, because inverse village size is the variable of interest, the most important variation comes from comparing small villages (e.g., 200 households) with very small villages (e.g., 50 households), while the few large villages (>1000 households) play very little role. That is, our analysis is not based on comparing urban areas with rural areas.[6] Third, no obvious geographic correlates with village size are discernible from GIS plots of village size in our area of study.

[5]More details of the funds and program are presented in Kaboski and Townsend (2007).

[6]Thus, any populist bias toward rural areas and against Bangkok is not likely to contaminate our analysis.

2.2 Reduced Form Impacts

In the analysis that follows we view the relationships between outcome measures and village size as driven by the program itself, and not the result of pre-existing differences in levels or trends. Indeed, this is established in depth by Kaboski and Townsend (2007), which presents reduced form regressions. We use a first-stage regression to predict village fund credit of household n in year t, $VFCR_{n,t}$:

$$\begin{aligned} VFCR_{n,t} &= \sum_{j=5,6} \alpha_{1,VFCR,j} \frac{1}{\#\text{ HHs in village}_{v,t}} \mathcal{I}_{t=j} + \alpha_{2,VFCR} \frac{1}{\#\text{ HHs in village}_{v,t}} \qquad (1) \\ &\quad + \boldsymbol{\gamma}_{VFCR} \mathbf{X}_{nt} + \theta_{VFCR,t} + \theta_{VFCR,n} + \varepsilon_{Z,nt} \end{aligned}$$

Here the crucial instrument is inverse village size in the post-intervention years (captured by the indicator function $\mathcal{I}_{t=j}$), X_{nt} is a vector of demographic household controls, and $\theta_{VFCR,t}$ and $\theta_{VFCR,n}$ are time and household specific-fixed effects respectively. Second stage outcome equations in levels and differences are of the form:

$$\begin{aligned} Z_{nt} &= \alpha_{1,Z} VFCR_{n,t} + \alpha_{2,\Delta Z} \frac{1,000,000}{\#\text{ HHs in village}_{v,t}} \qquad (2) \\ &\quad + \boldsymbol{\gamma}_{Z} \mathbf{X}_{nt} + \theta_{Z,t} + \theta_{Z,n} + \varepsilon_{Z,nt} \\ \Delta Z_{nt} &= \alpha_{1,\Delta Z} VFCR_{n,t} + \alpha_{2,\Delta Z} \frac{1,000,000}{\#\text{ HHs in village}_{v,t}} \qquad (3) \\ &\quad + \boldsymbol{\gamma}_{\Delta Z} \mathbf{X}_{nt} + \theta_{\Delta Z,t} + \theta_{\Delta Z,n} + \varepsilon_{\Delta Z,nt} \end{aligned}$$

Here Z_{nt} represents an outcome variable of interest for household n in year t**,** $\mathcal{I}_{t=j}$ is an indicator function used to capture the post-intervention years, X_{nt} is a vector of demographic household controls, and $\theta_{Z,t}$ and $\theta_{Z,n}$ are time and household specific-fixed effects respectively. Although there is heterogeneity across households and non-linearity in the impact of credit, $\hat{\alpha}_1$ captures (a linear approximation of) the relationship between the average impact of a dollar of credit on the outcome of Z_{nt}. The estimates of α_2 capture any relationship between village size and credit or outcome variables that exists before the program years. Using these and other specifications, we analyze a wide range of outcome variables Z_{nt}, including credit disaggregated by source and stated use, interest rates, income, investment, savings, lending, and assets.[7] The $\hat{\alpha}_2$ estimates are only significant in 2 of the 37 regressions, which is approximately the rate of Type 1 errors expected with a 5 percent significance level.[8]

The regressions produce several interesting "impact" estimates $\hat{\alpha}_{1,Z}$ as reported in detail in our companion paper, Kaboski and Townsend (2007).[9] Using village fund credit, total short-term credit and average interest rate on short-term credit as outcome variables, Z, shows how the injection expanded credit. First, the program expanded village fund credit roughly one for one, with the coefficient $\hat{\alpha}_{1,Z}$ close to one for village fund credit. Second, total credit overall appears to have had a similar expansion, and there is no evidence of crowding out in the credit market. Finally, the expansion did not occur through a reduction in interest rates. Indeed the $\hat{\alpha}_{1,Z}$ is positive, though small.

Household consumption was obviously and significantly effected by the program, with a $\hat{\alpha}_{1,Z}$ point estimate near one. The growth in consumption was driven by non-durable consumption and services, rather than durable goods. While the frequency of agricultural investments did increase mildly, total investment showed no significant response to the program. Default rates decreased in the year of the program, and then increased mildly in the second year, but remained less than 15 percent of loans. Asset levels (including savings) declined in response to the program, while income growth increased weakly.[10]

Together, these results are puzzling. In a perfect credit, permanent income model, unsubsidized credit should have no effect, while subsidized credit would simply have an

[7] An additional specification that we consider included a geographic control variable that capturing the average size of neighboring villages.

[8] Over the full sample, smaller villages were associated with *more* borrowing from commercial banks and lower consumption expenditures on fuel.

[9] The sample in Kaboski and Townsend (2007) varies slightly from the sample in this paper. Here we necessarily exclude 79 households who realized negative income (net of expenses) in any year.. To avoid confusion, we do not report the actual Kaboski and Townsend (2007) estimates here.

[10] Wage income also increased in response to the shock, which is a focus of Kaboski and Townsend (2007). The increase is quite small relative to the increase in consumption, however, and so this has little promise in explaining the puzzles. We abstract from general equilibrium effects on the wage and interest rate in the model we present.

income effect. If credit did not need to be repaid, this income effect would be bounded above by the amount of credit injected. Yet repayment rates were high enough that credit in the second year was equal to that in the first. Even if credit were not repaid, an income effect would produce at most a coefficient of the market interest rate (about 0.06), i.e., the household would keep the principle of the one-time wealth shock and consume the interest. The fact that households appear to have simply consumed the value of the fund. coefficient is therefore puzzling. Given the positive level of observed investment, the lack of a response to investment might point to well-functioning credit markets, but the large response of credit and consumption indicate the opposite. Thus, the coefficients overall require a theoretical and quantitative explanation.

2.3 Underlying Environment

Growth, savings/credit, default, and investment are key features in the Thai villages during the pre-intervention period (as well as afterward). Households income growth averages 7 percent over the panel, but both income levels and growth rates are stochastic. Savings and credit are important buffers against income shocks, but credit is limited. Transitory and permanent income shocks are neither fully insured nor fully smoothed, and Karaivanov and Townsend (2008) conclude that savings and borrowing models and savings only models fit the data better than alternative mechanism design models. High income households appear to have access to greater credit. Among borrowing households, regressions of log short-term credit on log current income yield a coefficient of 0.32 (std. err.=0.02).

Related, default occurs in equilibrium, and appears to one way of smoothing against shocks. In any given year, 11 percent of households are over three months behind in their payments on short-term (less than on year) debts. Default is negatively related to current income, but household consumption is substantial during periods of default, averaging 140 percent of current income, and positively related to income. Using only years of default, regressions of log consumption on log income yield a coefficient of regression of 0.44 (std. error=0.04).

Finally, investment plays an important role in the data, averaging 9 percent of income, but is lumpy. On average only 12 percent of households invest in any given year. Investment is large in years when investment occurs and highly skewed with a mean of 77 percent of total income and a median of 14 percent. Both the size of investment and the frequency of investment are positively associated with income, but high income households still invest infrequently. The top income tercile invest more often than poorer households but still only 22 percent of the time. Related, investment is not concentrated among the same households each year. On average 12 percent of households invest in each year. If this percent were independent across years, one would predict that 47 percent of households would invest at least once over the five years of pre-intervention data. This is quite close to the 42 percent that is observed.

The next section develops a model broadly consistent with this underlying environment.

3 Model

We address these key features of the data, by developing a model of a household facing permanent and transitory income shocks, that makes decisions about consumption, low yield liquid savings, high yield illiquid investment and default. In order to allow for growth, tractability requires that we make strong functional form assumptions. In particular, the problem is written so that all constraints are linear in permanent income, so that the value function and policy functions can all be normalized by permament income to attain a stationary, recursive problem.

3.1 Sequential Problem

In period 0, the household begins with a potential investment project of size I_0^*, a permanent component of income P_0, and liquid wealth L_0. At $t+1$,Liquid wealth L_{t+1} includes the principal and interest on liquid savings from the previous period $(1+r)\,S_t$ and current realized income Y_{t+1}:

$$L_{t+1} \equiv Y_{t+1} + S_t(1+r) \tag{4}$$

Following the literature on precautionary savings (e.g., Zeldes, 1989, Carroll, 1997, Gourinchas and Parker, 2001), current income Y_{t+1} consists of a permanent component of income P_{t+1} and a transitory one-period shock, U_{t+1}, additive in logs:

$$Y_{t+1} \equiv P_{t+1}U_{t+1} \tag{5}$$

We follow the same literature in modeling an exogenous component of permanent income that follows a random walk (again in logs) based on shock N_t with drift G, but our innovation is to allow for permanent income to also be increased endogenously through investment. Investment is indivisible – the household makes a choice $D_{I,t} \in \{0, 1\}$ of whether to undertake a lumpy investment project of size I_t^* or to not invest at all. In sum,

$$P_{t+1} = P_t G N_{t+1} + R D_{I,t} I_t^* \tag{6}$$

Investment is also illiquid and irreversible, but again it increases permanent income, at a rate R, higher than the interest rate on liquid savings, r, and sufficiently high to induce investment for households with high liquidity. Having investment increase the permanent component of future income simplifies the model by allowing us to track only P_t rather than multiple potential capital stocks, but it precludes some issues of investment "portfolio" decisions and risk diversification, which we believe are not of first-order interest.[11]

Project size is stochastic, governed by an exogenous shock i_t^* and proportional to permanent income:

$$I_t^* = i_t^* P_t \tag{7}$$

[11] Although investment I_t increases the permanent component of income in the following period P_{t+1} by a deterministic amount RI_t, it increases next period liquidity L_t stochastically, since income Y_t is the *product* of the permanent component and a transitory shock. In contrast, the net yield on savings increases next period liquidity by a deterministic rS_t. Furthermore, investment increases future permanent income P_{t+2} stochastically, since current permanent income is multiplied by a stochastic shock in contributing to future permanent income. See equations (4)-(6) above.

We assume that investment opportunities I_t^* increase with permanent income P_t. This is consistent with the empirical literature, where investment is typically scaled by size, i.e., large firms invest higher amounts. It also ensures that high permanent income alone will not automatically eliminate credit constraints and allow for investment every period. We do not observe this in the data, as discussed in the previous section.

Liquid savings can be negative, but borrowing is bounded below by a limit which is a multiple $\underline{s}$ of permanent income. That is, when $\underline{s}$ is negative borrowing is allowed, and this is the key parameter that we calibrate to the intervention:

$$S_t \geq \underline{s} P_t \tag{8}$$

For the purposes of this partial equilibrium analysis, this borrowing constraint is exogenous. It is not a natural borrowing constraint as in Aiyagari (1994) and therefore somewhat ad hoc, but such a constraint can arise endogenously in models with limited commitment (see Wright, 2002) or where lenders have rights to garnish a fraction of future wages (e.g., Lochner and Monge-Naranjo, 2008). Most importantly, it allows for default, which is observed in the data and of central interest to microfinance interventions.

The household's problem is to maximize expected discounted utility by choosing a sequence of consumption $C_t > 0$, savings S_t, and decisions $D_{I,t} \in \{0, 1\}$ of whether or not to investment:

$$\begin{aligned} V(L_0, I_0^*, P_0; \underline{s}) &= \max_{\substack{\{C_t>0\} \\ \{S_{t+1}\} \\ \{D_{I,t}\}}} E_0 \left[\sum\nolimits_{t=0}^{\infty} \beta^t \frac{C_t^{1-\rho}}{1-\rho} \right] \\ & \quad \text{s.t. eq. (4), (5), (6), (7), (8), and} && (9) \\ C_t + S_t + D_{I,t} I_t^* &\leq L_t && (10) \end{aligned}$$

The expectation is taken over sequences of permanent income shocks N_t, transitory income shocks U_t, and investment size shocks i_t^*. These shocks are each i.i.d. and orthogonal to one another:

- N_t is random walk shock to permanent income. $\ln N_t \sim N(0, \sigma_N^2)$.
- U_t is a temporary (one period) income shock. $u_t \equiv \ln U_t \sim N(0, \sigma_u^2)$.
- i_t^* is project size (relative to permanent income). $\ln i_t^* \sim N(\mu_i, \sigma_i^2)$

If $\underline{s} < 0$, an agent with debt, i.e., $S_{t-1} < 0$, and a sufficiently low income shock may need to default. That is, no positive value of consumption would satisfy equations (10) and (8). Essentially, given (8), a bad enough shock to permanent income (i.e., a low N_t) can produce a "margin call" on credit that exceeds current liquidity. In this case, we assume default allows for a minimum consumption level that is proportional to permanent income ($\underline{c}P_t$).

Defining the default indicator, $D_{def,t} \in \{0, 1\}$, this condition for default is expressed:

$$D_{def,t} = \begin{array}{l} 1, \text{ if } (\underline{s} + \underline{c})\, P_t < L_t \\ 0, \text{otherwise} \end{array} \tag{11}$$

and the defaulting household's policy for the period becomes:

$$\begin{aligned} C_t &= \underline{c}P_t \\ S_t &= \underline{s}P_t \\ D_{I,t} &= 0 \end{aligned}$$

This completes the model. The above modeling assumptions are strong, but motivated by the data. While these assumptions are not without costs, they do have analytical benefits. First, the model is simple and has limited heterogeneity, but consequently has a low dimension, tractable state space $\{L, I^*, P\}$ and parameter space $\{\sigma_N, \sigma_u, \mu_i, \sigma_i, G, r, R, \rho, \beta, \underline{s}, \underline{c}\}$. Hence, the role of each parameter and state can be more easily understood. The linearity of the constraints in P_t reduces the dimensionality of the state space to two, which allows for graphical representation of policy functions (in Section 5.2). The next subsection derives the normalized, recursive representation.[12]

[12] Conditions for the equivalence of the recursive and sequential problems and existence of the steady state are straightforward extensions of conditions given in Alvarez and Stokey (1998) and Carroll (2004). In particular, for $\rho < 1$, G and $RE\,[i^*]$ must be sufficiently bounded.

3.2 Normalized and Recursive Problem

Here, we have explicitly emphasized the value function's dependence on $\underline{s}$, since this will be the parameter of most interest in considering the microfinance intervention in Section 5. We drop this emphasis in the simplifying math that follows. Using lower case variables to indicate variables normalized[13] by permanent income, the recursive problem becomes:

$$
\begin{aligned}
V(L, I^*, P) &\equiv P^{1-\rho} v(l, i^*) \\
v(l, i^*) &= \max_{c, s', d_I} \frac{c^{1-\rho}}{1-\rho} + \beta E\left[(p')^{1-\rho} v\left(l', i^{*\prime}\right) \right] && (12) \\
& \quad s.t. \\
\lambda &: c + s + d_I i^* \leq l \qquad \text{from (10)} && (13) \\
\phi &: s \geq \underline{s} \qquad \text{from (8)} && (14) \\
p' &= GN' + R d_I i^* \qquad \text{from (6)} && (15) \\
l' &= y' + \frac{s(1+r)}{p'} \qquad \text{from (4)} && (16) \\
y' &= U' \qquad \text{from (5)} && (17)
\end{aligned}
$$

We further simplify by substituting in for l' and y' into the continuation value using (16) and (17), and substituting out s using the liquidity budget constraint (**??**), which will hold with equality, to yield:

$$
\begin{aligned}
v(l, i^*) &= \max_{c,\, d_I} \frac{c^{1-\rho}}{1-\rho} \\
& \quad + \beta E\left[(p')^{1-\rho} v\left(U' + \frac{(1+r)(l - c - d_I i^*)}{p'}, i^{*\prime} \right) \right] && (18) \\
& \quad s.t. \\
\phi &: (l - c - d_I i^*) \geq \underline{s} && (19) \\
p' &= GN' + R d_I i^* && (20)
\end{aligned}
$$

[13] Here the decision whether to invest d_i is not a normalized variable and is in fact identical to D_i in the sequential problem. We have denoted it in lower case to emphasize that it will depend only on the normalized states l and i^*.

The normalized form of the problem has two advantages. First, it lowers the dimensionality of the state variable. Second, it allows the problem to have a steady state solution. Using $*$ to signify optimal decision rules, the necessary conditions for optimal consumption c_* and investment decisions d_{I*} are:[14]

$$(c_*)^{-\rho} = \beta(1+r)E\left[(p')^{-\rho}\frac{\partial v}{\partial l}(U' + \frac{(1+r)(l - c_* - d_{I*}i^*)}{p'}, i^{*\prime})\right] + \phi \tag{21}$$

$$\begin{aligned} \frac{c_*^{1-\rho}}{1-\rho} + \beta E\left[(p')^{1-\rho} v\left(U' + \frac{(1+r)(l - c_* - d_{I*}i^*)}{p'}, i^{*\prime}\right)\right] &\geq \\ \frac{c_{**}^{1-\rho}}{1-\rho} + \beta E\left[(p')^{1-\rho} v\left(U' + \frac{(1+r)\left[l - c_{**} - (1-d_{I*})i^*\right]}{p'}, i^{*\prime}\right)\right] & \end{aligned} \tag{22}$$

where c_{**} indicates the optimal consumption given the alternative investment decision (i.e., c_{**} satisfies the analog to (21) for $1 - d_{I*}$). The constraint ϕ is only non-zero when (19) binds, i.e., $c_* = l - \underline{s} - d_{I*}i^*$.

In practice, the value function and optimal policy functions must be solved numerically, and indeed the indivisible investment decision complicates the computation.[15]

Figure 1 presents a three-dimensional graph of a computed value function. The flat portion at very low levels of liquidity l comes from the minimum consumption and default option. The dark line highlights a groove going through the middle of the value function surfaces along the critical values at which households first decide to invest in the lumpy project. These threshold levels of liquidity are increasing in the size of the project. The slope of the value function with respect to l increases at this point because the marginal utility of consumption increases at the point of investment.[16] Consumption actually *falls* as liquidity increases beyond this threshold.

Figure 2, panel A illustrates this more clearly by showing a cross-section of the optimal consumption policy function at a given value of i^*. At low values of liquidity, no investment

[14] Although the value function is kinked, it is differentiable almost everywhere, and the smooth expectation removes any kink in the continuation value.

[15] Details of the computational approach and codes are available from the authors upon request.

[16] Given the convex kink in the value function, households at or near the kink would benefit from lotteries, which we rule out.

is made, households consume as much as possible given the borrowing constraint, and hence the borrowing constraint holds with equality. At higher liquidity levels, this constraint is no longer binding as savings levels s exceed the lower bound $\underline{s}$. At some crucial level of liquidity l_*, the household chooses to invest in the lumpy project, at which point consumption falls and the marginal propensity to consume out of additional liquidity increases. Although not pictured, for some parameter values (e.g., very high R), the borrowing constraint can again hold with equality, and marginal increases in liquidity are used for purely for consumption.[17]

Panel B shows the effect of a surprise permanent decrease in $\underline{s}$ on the optimal consumption policy for the same given value of i^*. Consumption increases for liquidity levels in every region, except for the region that is induced into investing by more access to borrowing.

An additional interesting prediction of the model is that for a given level of borrowing ($s_t < 0$), a household that invests ($d_{I,t} = 1$) has a lower probability of default. Conditional on investing, the default probability is further decreasing in the size of investment. Thus, other things equal borrowing to invest leads to less default than borrowing to consume because investment increases further income and therefore ability to repay. The maximum amount of debt that can be carried over into next period (i.e., $-\underline{s}P_t$) is proportionate to permanent income. Because investment increases permanent income, it increases the borrowing limit next period, and therefore reduces the probability of a "margin call" on outstanding debt.

One can see this formally by substituting the definitions of liquidity (4) and income (5), and the law of motion for permanent income (6) into the condition for default (11) to

[17] Using a bufferstock model, Zeldes (1989) derived reduced form equations for consumption growth, and found that consumption growth was significantly related to current income, but only for low wealth households, interpreted as evidence of credit constraints. We run similar consumption growth equations that also contain investment as an explanatory variable:

$$\ln C_{n,t+1}/C_{n,t} = X_{n,t}\beta_1 + \beta_2 Y_{n,t} + \beta_3 I_{n,t} + \varepsilon_{n,t}$$

and for the low wealth sample, we find significant estimates $\hat{\beta}_2 < 0$ and $\hat{\beta}_3 > 0$, which is consistent with the prediction of investment lowering current consumption (thereby raising future consumption growth).

yield:

$$E\left(D_{def,t+1}\right) = \Pr\left[U_{t+1} < (\underline{s} + \underline{c}) - \frac{S_t}{(P_t N_{t+1} G + R D_{I,t} I_t^*)}\right] \tag{23}$$

Since S_t is negative and R is positive, the right-hand side of the inequality is decreasing in both $D_{I,t}$ and I_t^*. Since both N_{t+1} and U_{t+1} are independent of investment, the probability is therefore decreasing in $D_{I,t}$ and I_t^*.

4 Estimation

This section addresses the estimation approach and then the data used. The model is quite parsimonious with a total of 11 parameters. Due to poor identification, we calibrate the return on investment parameter, R, using a separate data source. We then estimate the remaining parameters, $\theta = \{r, G, \sigma_N, \sigma_u, \underline{s}, \underline{c}, \beta, \rho, \mu_i, \sigma_i\}$ via GMM using the liming optimal weighting matrix. This estimation is performed using five years (1997-2001) of pre-intervention data, so that $t = 1$ corresponds to the year 1997.

4.1 Data

The data come from the Townsend Thai data project, an ongoing panel dataset of a stratified, clustered, random sample of institutions (256 in 2002), households (960 each year, 720 with complete data in the pre-experiment balanced panel used for estimation, and 719 and 705 in 2002 and 2003, respectively, which are used only for prediction), and key informants for the village (64, one in each village). The data are collected from sixty-four villages in four provinces: Buriram and Srisaket in the Northeast region, and Lopburi and Chachoengsao in the Central region. The components used in this study include detailed data from households and household businesses on their consumption, income, investment, credit, liquid assets and the interest income from these assets, as well as village population data from the village key informants. All data has been deflated using the Thai consumer price index to the middle of the pre-experiment data, 1999.

The measure of household consumption we use (denoted $\tilde{C}_{n,t}$ for household n at time t) is calculated using detailed data on monthly expenditure data for thirteen key items, and scaled up using weights derived from the Thai Socioeconomic Survey.[18] In addition, we include household durables in consumption, though durables play no role in the observed increases in consumption. The measure of investment ($\tilde{I}_{n,t}$) we use is total farm and business investments.

We impute default each year for households who report one or more loans due in the previous 15 months that are outstanding at least three months. Note that (1) this includes all loans, and not just short-term, since any (non-voluntary) default indicates a lack of available liquidity, and (2) due dates are based on the original terms of the loan, since changes in duration are generally a result of default.[19] This only approximates default in the model, and it may underestimate default because of underreporting, but overestimate default as defined in the model or to the extent that late loans are eventually repaid.

The income measure we use (denoted $\tilde{Y}_{n,t}$) includes all agricultural, wage, business and financial income (net of agricultural and business expenses) but excludes interest income on liquid assets such as savings deposits. Our savings measure ($S_{n,t}$) includes not only savings deposits in formal and semi-formal financial institutions, but also the value of rice holdings in the household. Cash holdings are unfortunately not available. The measure of liquid credit ($CR_{n,t}$) is short-term credit with loan durations of one year or less. The measurement of interest income on liquid savings ($EARNED\_INT_{n,t}$) is interest income in year t on savings in formal and semi-formal institutions. The interest owed on credit ($OWED\_INT_{n,t}$) is the reported interest owed on short-term credit.

While the data is high quality and detailed, measurement error is an important concern. Net income measures may complicated when expenditures and corresponding income do not coincide in the same year, for example. If income is measured with error, the amount

[18] The tildes represent raw data which will be normalized in Section 4.3.1.

[19] According to this definition, default probability is about 19 percent, but alternative definitions can produce different results. The probability for short-term loan alone is just 12 percent, for example. Labeling all loans from non-family sources that have no duration data whatsoever as in default yields a default probability of 23 percent. Our results for consumption and default hold for the higher rates of default.

of true income fluctuations will be overstated in the data, and household decisions may appear to be less closely tied to transitory income shocks, and hence credit constraints may not appear to be important. Consumption and investment may also suffer from measurement error, but classical measurement error will just add additional variation to these endogenous variables will not effect the moments, only the weighting matrix. A major source of measurement error on interest is that savings and borrowing may fluctuate within the year, so that both earned and paid interest may not accurately reflect interest on the measured end of year stocks. Such measurement error will assist in the estimation.

Table 1 presents key summary statistics for the data.

4.1.1 Adjusting the Data for Demographic and Cyclical Variation

The model is of infinitely lived dynasties that are heterogeneous only in their liquidity, permanent income, and potential investment. The data, however, contain important variation in household composition across time and households. The data also contain business cycle variation which is not included in the model. Ignoring either of these sources of variation would be problematic. For household composition, to the extent that changes in household composition are predictable, the variance in income changes may not be capturing uncertainty but also predictable changes in household composition. Likewise, consumption variation may not be capturing household responses to income shocks, but predictable responses to changes in household composition. Failure to account for this would likely exaggerate both the size of income shocks and the response of household consumption to these shocks. In the data, the business cycle (notably the financial crisis in 1997 and subsequent recovery) also plays an important role in household behavior, investment and savings behavior in particular. Although our post-program analysis will focus on the across-village differential impacts of the village fund program, and not merely the time-changes, we do not want to confound the impacts with business cycle movements.

We therefore follow Gourinchas and Parker (2002) in removing the business cycle and household composition variation from the data. In particular, we run linear regressions of log income, log consumption, and liquidity over income. (We do not take logs of liquidity, since it takes both positive and negative values, but instead normalize by income so that high values do not carry disproportionate weight.)[20] The estimated equations are:

$$\begin{aligned}
\ln \tilde{Y}_{nt} &= \boldsymbol{\gamma}_Y \mathbf{X}_{nt} + \theta_{Y,jt} + e_{Y,nt} \\
\tilde{L}_{nt}/\tilde{Y}_{nt} &= \boldsymbol{\gamma}_L \mathbf{X}_{nt} + \theta_{L,jt} + e_{L,nt} \\
\ln \tilde{C}_{nt} &= \boldsymbol{\gamma}_C \mathbf{X}_{nt} + \theta_{C},_{jt} + e_{C,nt} \\
\ln \tilde{D}_{nt} &= \boldsymbol{\gamma}_D \mathbf{X}_{nt} + \theta_{D,jt} + e_{D,nt}
\end{aligned}$$

where X_{nt} is a vector of household composition variables (i.e., number of adult males, number of adult females, number of children, male head of household dummy, linear and squared terms of age of head of household, years education of head of household, and a household-specific fixed effect) for household n at time t and θ_{jt} is a time t-specific effect that varies by region j and captures the business cycle. Unfortunately, these time-specific effects cannot be extrapolated for the post-program data, so we rely on across village, within-year variation to evaluate the model's predictions. These regressions are run using only the pre-program data, 1997-2001. The R^2 values for the four regressions are small: 0.06, 0.001, 0.22, and 0.07, respectively.

For the full sample, 1997-2003, we construct the adjusted data for a household with mean values of the explanatory variables ($\overline{X}$ and $\overline{\theta}_j$) using the estimated coefficients and residuals:

$$\begin{aligned}
\ln Y_{nt} &= \hat{\boldsymbol{\gamma}}_Y \overline{\mathbf{X}} + \overline{\theta}_{Yj} + g(t-1999) + \hat{e}_{Y,nt} \\
L_{nt}/Y_{nt} &= \hat{\boldsymbol{\gamma}}_L \overline{\mathbf{X}} + \overline{\theta}_{Lj} + \hat{e}_{L,nt} \\
\ln C_{nt} &= \hat{\boldsymbol{\gamma}}_C \overline{\mathbf{X}} + \overline{\theta}_{Cj} + \hat{e}_{C,nt} \\
D_{nt} &= \hat{\boldsymbol{\gamma}}_D \overline{\mathbf{X}} + \overline{\theta}_{Dj} + \hat{e}_{D,nt}
\end{aligned}$$

[20] As noted before, 79 of the original 960 households realized negative net income at some point in the pre-intervention sample. The model yields only positive income, and so these households were dropped.

where g is the average growth rate of income in the pre-program data. Next, we use a multiplicative scaling term to ensure that average income, liquidity ratios, consumption, and default are equal in the raw and adjusted data. Finally, we construct investment data I_{nt} by multiplying the actual measured values of i_{nt} (i.e., $i_{nt} \equiv \tilde{I}_{nt}/\tilde{Y}_{nt}$) by the newly constructed income data Y_{nt}.

4.2 Returns on Investment

In principle, income growth and investment data should tell us something about the return on investment, R. In practice, however, the parameter cannot be well estimated because investment data itself is endogenous to current income, and also because investment occurs relatively infrequently. We instead use data on physical assets rather than investment, and we calibrate R to match cross-sectional relationship between assets and income.

To seperate the effect of assets and labor quality on income, we assume that all human capital investments are made prior to investments in physical assets. Let $t - J$, indicate the first year of investing in physical assets. That is, substituting the law of motion for permanent income, equation (6), J times recursively into the definition of actual income, equation (5), yields:

$$Y_t = \underbrace{\left[P_{t-J} G^J \prod\nolimits_{j=1}^{J} N_{t+1-j}\right] U_t}_{\text{income of investment prior to } t-J} + \underbrace{R \left[\sum_{j=1}^{J} I_{t-j} G^{j-1} \prod\nolimits_{k=1}^{j} N_{t+1-j}\right] U_t}_{\text{income from investment after } t-J}$$

The first term captures income from the early human capital investments, which we measure by imputing wage income based on household characteristics (sex, education, region). The second term involves the return R multiplied by the some of the past J years of investments (weighted by the deterministic and random components of growth.) We measure this term using current physical assets. That is, R is the calibrated using the following operational formula:

$$\varepsilon_{R,t} = Y_t - \text{imputed labor income}_t - R\,(\text{physical assets}_t)$$

We have the additional issue of how to deal with the value of housing and unused land. Neither source of assets contributes to Y_t, so we would ideally exclude them from the stock of assets.[21] Using data on the (a) value of the home, (b) value of the plot of land including the home, and (c) the value of unused or community use land, we construct three variants of physical assets.

We use a separate data set, the Townsend Thai Monthly Survey, to calibrate this return. The data is obtained from different villages, but the same overall survey area, and the monthly has the advantage of including wage data used to impute the labor income portion of total income.

We us a procedure which is analogous to GMM. We choose R to set the average $\varepsilon_{R,t}$ to zero in the sample of households. The baseline value (which excludes categories (a)-(c) from assets) yields $R = 0.11$, while including (c), or (b) and (c), yield $R = 0.08$ and $R = 0.04$, respectively. If we choose R to solve $\varepsilon_{R,t} = 0$ for each household, then the median R values are identical to our estimates. Not surprisingly, R substantially varies across households, however. This is likely due in part because permament shock histories and current transitory shocks differ across households, but also in part because households face different ex ante returns to investment.

4.3 Method of Simulated Moments

In estimating, we introduce multiplicative measurement error in income which we assume is log normally distributed with log mean 0 and standard deviation σ_E. We therefore have twelve remaining parameters $\theta = \{r, G, \sigma_N, \sigma_u, \sigma_E, \underline{c}, \beta, \rho, \mu_i, \sigma_i, \underline{s}\}$, which are estimated using a Method of Simulated Moments.The model parameters are identified jointly by the full set of moments, but I intuitively discuss the specific moments that are particularly important for identifying each parameter.

[21] Our measure of Y_t does not include imputed owner occupied rent.

The first two types of moments help identify the return to liquid savings, r :

$$\begin{aligned}\varepsilon_s(X,r) &= EARNED\_INT_t - rS_{t-1}\\ \varepsilon_{cr}(X,r) &= OWED\_INT_t - rCR_{t-1}\end{aligned}$$

In ε_s, S_{t-1} is liquid savings in the previous year, while $EARNED\_INT_t$ is interest income received on this savings. Likewise, in ε_{cr}, CR is outstanding short-term credit in the previous year, and $OWED\_INT$ is the subsequent interest owed on this short-term credit in the following year.[22]

The remaining moments require solving the policy functions for consumption, $C(L_t, P_t, I_t^*; \theta) = P_t c(l_t, i_t^*; \theta)$, investment decisions, $D_I(L_t, P_t, I_t^*; \theta) = d_I(l_t, i_t^*; \theta)$, and default decisions, $D_{def}(L_t, P_t; \theta) = d_{def}(l_t; \theta)$,where we have explicitly denoted their dependence on the parameter set θ. We observe data on decisions, C_t, I_t, $D_{def,t}$, and states L_t and Y_t. Our strategy is to use these policy functions to define deviations of actual variables (policy decisions and income growth) from the corresponding expectations of these variables conditional on L_t and Y_t.[23] By the Law of Iterated Expectations, these deviations are zero in expectation and therefore valid moment conditions. The details of calculating the conditional expectations (over the shocks U_t, N_t, and i_t^*) are in Appendix A.

The income growth moments help to identify the income process parameters and are derived from the definition of income and the law of motion for permanent income, equations (5) and (6).[24] Average income growth helps identify the drift component of growth income growth, G:

$$\varepsilon_g(L_t, Y_t, Y_{t+1}; \theta) = \ln(Y_{t+1}/Y_t) - E[\ln(Y_{t+1}/Y_t) | L_t, Y_t]$$

The variance of income growth over different horizons (k =1...3-year growth rates, respectively) helps identify standard deviation of transitory and permanent income shocks, σ_u

[22] In the data there are many low interest loans, and the average difference between households interest rates on short term borrowing and saving is small, just 2 percent.

[23] Since L_t requires the previous years savings S_{t-1}, these moments are not available in the first year.

[24] Carroll and Samwick (1997) provide techniques for estimating the income process parameters G, σ_N, and σ_u without solving the policy function. These techniques cannot be directly applied in our case, however, since income is depends on endogenous investment decisions.

and σ_N, since transitory income shocks add the same amount of variance to income growth regardless of k, wheras the variance contributed by permanent income shocks increases with k. Still, it is difficult to distinguish transitory income shocks from measurement error from income growth variance alonge. The deviations are defined as:

$$\begin{aligned}
\varepsilon_{v,k}(L_t, Y_t, Y_{t+j}; \theta) &= \begin{bmatrix} \ln\left(Y_{t+k}/Y_t\right) \\ -E\left[\ln\left(Y_{t+k}/Y_t\right)|L_t, Y_t\right] \end{bmatrix}^2 \\
&\quad -E\left[\begin{bmatrix} \ln\left(Y_{t+k}/Y_t\right) \\ -E\left[\ln\left(Y_{t+k}/Y_t\right)|L_t, Y_t\right] \end{bmatrix}^2 \middle| L_t, Y_t\right] \\
\text{for } k &= 1, 2, 3
\end{aligned}$$

we identify minimum consumption, $\underline{c}$; the project size distribution parameters, μ_i and σ_i; and the preference parameters β and ρ using moments on consumption decisions, investment decisions, and the size of investments. We focus on deviations in log consumption, investment decisions, and log investments (when investments are made):

$$\begin{aligned}
\varepsilon_C(C_t, L_t, Y_t; \theta) &= \ln C_t - E\left[\ln C_t | L_t, Y_t\right] \\
\varepsilon_D(D_{I,t}, L_t, Y_t; \theta) &= \ln D_{I,t} - E\left[D_{I,t} | L_t, Y_t\right] \\
\varepsilon_I(D_{I,t}, I_t, L_t, Y_t; \theta) &= D_{I,t} \ln I_t - E[D_{I,t} \ln I_t^* | L_t, Y_t]
\end{aligned}$$

We are left with essentially three moment conditions for five parameters, but gain additional moment conditions by realizing that since these deviations are conditional on income and liquidity, their interaction with functions of income and liquidity should also be zero in expectation. Intuitively, in expectation, the model should match average log consumption, probability of investing, and log investment across all income and liquidity levels, e.g., not overpredicting at low income or liquidity levels, while underpredicting at high levels. These moments play particular roles in identifying measurement error shocks σ_E and $\underline{c}$, in particular. If the data shows less response of these policy variables to income then predicted that could be due to a high level of measurement error in income. Similarly, high consumption at low levels of income and liquidity in the data would indicate a high level of minimum consumption $\underline{c}$.

Omitting the functional dependence of these deviations, we express below the nine valid moment conditions:

$$\begin{array}{ccc} E\left[\varepsilon_C\right]=0 & E\left[\varepsilon_D\right]=0 & E\left[\varepsilon_I\right]=0 \\ E[\varepsilon_C \ln Y_t]=0 & E[\varepsilon_D \ln Y_t]=0 & E[\varepsilon_I \ln Y_t]=0 \\ E[\varepsilon_C\left(L_t/Y_t\right)]=0 & E[\varepsilon_D\left(L_t/Y_t\right)]=0 & E[\varepsilon_I\left(L_t/Y_t\right)]=0 \end{array}$$

Finally, given $\underline{c}$, default decision moments are used to identify the borrowing constraint $\underline{s}$, which can be clearly seen from equation (11):

$$\varepsilon_{def}\left(L_t, Y_t, D_{def,t}\right)=D_{def,t}-E\left[D_{def,t}|Y_t, L_t\right]$$

In total, we have 16 moments to estimate 11 parameters.

4.4 Estimation Results

Table 2 presents the estimation results for the structural model. The interest rate $\hat{r}$ (0.054) is midway between the average rates on credit (0.073) and savings (0.035), and is quite similar to the six percent interest rate typically charged by village funds. The estimated discount factor $\hat{\beta}$ (0.915) and elasticity of substitution $\hat{\rho}$ (1.16) are within the range of usual values for bufferstock models. The estimated standard deviations of permanent $\hat{\sigma}_N$ (0.31) and transitory $\hat{\sigma}_U$ (0.42) income shocks are about twice those for wage earners in the United States (see Gourinchas and Parker, 2002), but reflect the higher level of income uncertainty of predominantly self-employed households in a rural, developing economy. In contrast, the standard deviation of measurement error $\hat{\sigma}_E$ (0.15) is much smaller than that of actual transitory income shocks, and is the only estimated parameter that is not significantly different from zero. The average log project size $\hat{\mu}_i$ greatly exceeds the average size of actual investments (i.e., $log\ I_t/Y_t$) in the data (1.47 vs. -1.96), and there is a greater variance in project size $\hat{\sigma}_i$ than in investments in the data (6.26 vs. 1.50). In the model, these difference between the average sizes of realized investment and potential projects stem from the fact that larger potential projects are much less likely to be undertaken. The estimated borrowing constraint parameter $\hat{\underline{s}}$ indicates that agents could borrow up to about 8 percent of their annual permanent income as short-term credit in the baseline period. (In the summary statistics of Table 1, credit averages about 20 percent of annual income, but liquid savings net of credit, the relevant measure, is actually positive and averages 9 percent of income.) The value of $\underline{\hat{c}}$ indicate consumption in default is roughly half of the permanent component of income.

Standard errors on the model are relatively small. Table 3 presents a matrix of partial derivatives of the 16 moments with respect to 11 parameter values in order to highlight which moments identify which parameters. The partial derivatives verify the intuition above. In particular, the interest rate r affects many moments, but is the only moment in the interest moments. While σ_N is important for the variance of two and three-year growth rates σ_U and σ_E are important for the variance of one-year growth rates, but σ_E also has a strong effect on the interaction of consumption and investment decisions with Y and L/Y. The utility function parameters β and ρ have the most important effect on consumption and investment moments. Also, while μ_i and σ_i also affect income growth variance, the investment probability and investment level moments also help identify them. Finally, both $\underline{s}$ and $\underline{c}$ affect default similarly, but have very opposite-signed effects on the interaction of investment and especially consumption decisions with measured income and liquidity ratios.

In terms of the fit, the model does well in reproducing average default probability, consumption, investment probability and investment levels, and indeed deviations are uncorrelated with log income or liquidity ratios. Still, we can easily rejects the overidentifying restrictions in the model, which tells us that the model is not the real world. The large J-statistic is driven by two sets of moments. First, the estimation rejects that the savings and borrowing rates are equal. Second, the model does poorly in replicating the volatility of the income growth process, yielding too little volatility.

We suspect this is the result of the income process and our statistical procedures failing to adequately capture cyclical effects of income growth, in particular the Thai financial crisis and recovery of 1997 and 1998 (survey years 1998 and 1999, respectively) Only *mean* time-varying volatility is extracted from the data using our regression techniques, but the crisis presumably affected the variance as well.[25] Excluding the crisis from the pre-sample is not possible, since it would leave us just one year of income growth to identify both transitory and permanent income shocks. An alternative estimation that uses only data from 2000 and 2001, except for 1999 data used to create two-year income growth variance moments produced estimates with wide standard errors that were not statistically different from the estimates above. The only economically significant difference was a much lower borrowing constraint ($\hat{\underline{s}} = -0.25$), which is consistent with an expansion of credit observed in the Thai villages.

So the result on the fit of the model are mixed. However, we view the model's ability to make policy predictions on the impact of credit as a stronger basis for evaluating its usefulness. We consider this in the next section.

5 Million Baht Fund Analysis

This section introduce the Million Baht fund intervention into the model, examines the model's predictions relative to the data, presents a normative evaluation of the program, and then presents alternative analyses allowed for by the structural model.

5.1 Relaxation of Borrowing Constraints

We incorporate the injection of credit into the model as a suprise decrease in $\underline{s}$.[26]. That is, for each village v – a subscript we now add to the notation – we calibrate the new, reduced constraint under the million baht fund intervention $\underline{s}_v^{mb}$ as the level for which our model

[25] We know from alternative estimation techniques that the model does poorly in matching year-to-year fluctuations in variables. In the estimation we pursue, we construct moments for consumption, investment, etc., that are based only on averages across the four years. For income growth volatitility, the moments necessarily have a year-specific component.

[26] Microfinance is often viewed as a lending technology innovation which is consistent with the reduction in $\underline{s}$. An alternative would be to model the expansion of credit through a decrease in the interest rate on borrowing, but recall that we did measure a decline in short-term interest rates in response to the program.

would predict one million baht of additional credit relative to the baseline at $\underline{s}$. We explain this mathematically b

below. Define first the expected borrowing of abelow. Define first the expected borrowing of a household n with the million baht fund intervention:

$$E\left[B_{n,t,v}^{mb}|L_{n,t},Y_{n,t};\underline{s}_v^{mb}\right] = E\left\{\mathcal{I}_{<0}\left[\begin{array}{c} L_t - C(L_t,P_t,I_t^*;\underline{s}_v^{mb}) \\ -D_I(L_t,P_t,I_t^*;\underline{s}_v^{mb})I_t^* \end{array}\right]|L_{n,t},Y_{n,t}\right\}$$

and in the baseline without the intervention:

$$E\left[B_{n,t,v}|L_{n,t},Y_{n,t};\underline{s}\right] = E\left\{\mathcal{I}_{<0}\left[\begin{array}{c} L_t - C(L_t,P_t,I_t^*;\underline{s}_v^{mb}) \\ -D_I(L_t,P_t,I_t^*;\underline{s}_v^{mb})I_t^* \end{array}\right]|L_{n,t},Y_{n,t}\right\}$$

where $\mathcal{I}_{<0}$ is shorthand notation for the indicator function that the bracketed expression is negative (i.e., borrowing and not savings). On average, village funds lent out 950,000 baht in the first year, so we choose $\underline{s}_v^{mb}$ so that we would have hypothetically predicted an additional 950,000 baht of borrowing in each village in the pre-intervention data [27]:

$$\frac{1}{\mathcal{N}}\sum_{n=1}^{\mathcal{N}}\left\{\begin{array}{c} E\left[B_{n,t,v}^{mb}|L_{n,t},Y_{n,t};\underline{s}_v^{mb}\right] \\ -E\left[B_{n,t,v}|L_{n,t},Y_{n,t};\underline{s}\right] \end{array}\right\} = \frac{950,000}{\#\text{ HHs in village}_v}$$

Here $\mathcal{N}$ represents the number of surveyed households in the pre-intervention data.

The resulting $\underline{s}_v^{mb}$ values average -0.28, with a standard error of 0.14, a minimum of -0.91 and a maximum of -0.09. Hence, for most villages, the post-program ability to borrow is substantial relative to the baseline ($\underline{s} = -0.08$), averaging about one-fifth of permanent income after the introduction of the program.[28]

5.2 Predictive Power

Using the calibrated values of borrowing limits, we evaluate the model's predictions for 2002 and 2003 (i.e., $t = 6$ and 7) on three dimensions: log consumption, probability of

[27] Since 1999 is the base year used, the 950,000 baht is deflated to 1999 values. Predicted results are similar if we use the one million baht which might have been predicted ex ante.

[28] These large changes are in line with the size of the intervention, however. In the smallest village, the ratio of program funds to village income in 2001 is 0.42. If half the households borrow, this would account for the 0.83 drop in $\underline{s}$.

investing, and log investment levels. Using the observed liquidity ($L_{n,5}$) and income data ($Y_{n,5}$) for year five (i.e., 2001), the last pre-intervention year, we draw series of $U_{n,t}$ and $i^*_{n,t}$ shocks from the estimated distributions, and simulate the model for 2002 and 2003. We do this 500 times, and combine the data with the actual pre-intervention data, in order to create 500 artificial datasets.

We then ask whether reduced-form regressions would produce similar impact estimates using simulated data as they would using the actual data, even though statistically the model is rejected, using the following reduced form regressions:

$$
\begin{aligned}
C_{n,t} &= \sum_{j=5,6} \alpha_{C,j} \frac{950,000}{\#\text{ HHs in village}_v} \mathcal{I}_{t=j} + \theta_{C,t} + e_{C,n,t} \\
D_{n,t} &= \sum_{j=5,6} \alpha_{D,j} \frac{950,000}{\#\text{ HHs in village}_v} \mathcal{I}_{t=j} + \theta_{D,t} + e_{D,n,t} \\
I_{n,t} &= \sum_{j=5,6} \alpha_{I,j} \frac{950,000}{\#\text{ HHs in village}_v} \mathcal{I}_{t=j} + \theta_{I,t} + e_{I,n,t} \\
DEF_{n,t} &= \sum_{j=5,6} \alpha_{DEF,j} \frac{950,000}{\#\text{ HHs in village}_v} \mathcal{I}_{t=j} + \theta_{DEF,t} + e_{DEF,n,t} \\
\ln\left(Y_{n,t}/Y_{n,t-1}\right) &= \sum_{j=5,6} \alpha_{\Delta \ln Y,j} \frac{950,000}{\#\text{ HHs in village}_v} \mathcal{I}_{t=j} + \theta_{\Delta \ln Y,t} + e_{\Delta \ln Y,n,t}
\end{aligned}
$$

Here $\hat{\alpha}_{C,j}$, $\hat{\alpha}_{D,j}$, $\hat{\alpha}_{I,j}$, $\alpha_{DEF,j}$, and $\alpha_{\Delta \ln Y,j}$ would be estimates of the year j impact of the program on consumption, investment probability, average investment, default probability, and log income growth, respectively. The equations only differ from the motivating regressions, equation (2), in three ways. First, we do not have a theory of actual borrowing from the village fund, so we have replaced predicted village fund credit with $\frac{950{,}000}{\#\text{ HHs in village}_v}$, the average injection per household. Second, we now show year-specific impacts, where the α's are the coefficients of interest. Third, the regressions above omit the household level controls and household fixed-effects, but recall that our data has already been purged of variation correlated with household level demographic data. We run these regressions on both the simulated and actual data and compare the estimates and standard errors. For the post-program years, the year fixed effects in the data will include the cyclical component, which the model is not intended to capture.

Table 3 compares the regression results of the model to the data, and shows that the model does generally quite well in replicating the results, particularly for consumption, investment probability, and investment.

The top panel presents the estimates from the actual data. These regressions yield the surprisingly high, and highly significant, estimates for consumption of 1.39 and 0.90 in the first year and second year, respectively. The estimate on investment probability is significant and positive, but only in the first year. For a village, with the average village fund credit per household of 9600, the point estimate of 6.3e-6 would translate into an increase in investment probability of six percentage points. Nonetheless, and perhaps surprising in a world without lumpy investment, the regressions find no significant impact on investment, and very large standard errors on the estimates. Similarly, the impact effects on default are insignificant. Finally, the impact of the program on log income growth is positive and significant, but only in the second year. Again, given the average village fund credit per household, this coefficient would into a ten percentage point higher growth rate in the second year.

The second panel of Table 4 presents the regressions using the simulated data. The first row shows the average (across 500 samples) estimated coefficient and the second row shows the average standard error on these estimates. The main point is that the estimates in the data are typical of the estimates the model produces for consumption, investment probability, and investment. In particular, the model yields a large and significant estimate of coefficient that is close to one in the first year, and a smaller though still large estimate in the second year. The standard errors are also quite similar to what is observed. The model also finds a comparably sized significant coefficient on the investment probabilities, although its average coefficients are more similar in both the first and second years, wheras the data show a steep drop off in the magnitude and significance after the first year.

The model's predictions for default and income volatility growth are less aligned with the data. For default, both the model and data show a marked and significant decrease in default in the first year, though the model's is much larger. While the data show a significant increase in default in the second year, the model produces no effect.[29] The data also shows a significant increase in income growth in the second year, whereas regressions from the model measure no impact on income growth. Perhaps, both of these shortcomings are results of the model's inability to fully capture year to year fluctuations in the *volatility* of the income growth process in the estimation.

The final panel shows formally that the estimates from the model are statistically similar to those in the data. It shows the fraction of simulations in which a Chow test on a sample with both the actual and simulated data for the post-program years finds a structural break at a 5 percent level of significance. Such occurences are quite rare. Except for investment levels, where outliers can drive results, structural breaks are found at a rate comparable to the level of significance.

One further note is that while the impact coefficients in the data are quite similar to those in the simulated structural model, they differ substantially from what would be predicted using reduced form regressions. That is, if we included income, liquidity, or credit as regressors in the reduced form equations, F-tests reject that the impacts coefficients of the program are the same as the coefficients on liquidity, income, or credit for the consumption, investment probability, and default regressions.

In sum, we measure large average effects on consumption and insignificant effects on investment, but the structural model helps us in quantitatively interpreting these impacts. First, these average coefficients mask a great deal of unobserved heterogeneity. Consider Figure 3 which shows the estimated policy function for consumption (normalized by permanent income) c as a function of (normalized) project size i^* and (normalized) liquidity l. The cliff-like drop in consumption running diagonally through the middle of the graph represents the threshold level of liquidity that induces investment. In the simulations, households in a village are distributed along this graph, and the distribution depends on the observables (Y and L), and stochastic draws of the shocks (i^* and U, since $P = \frac{Y}{U}$).

We have plotted examples of five potential households, all of whom could appear ex ante

[29] For the alternative definition of default, where all loans not relative loans with an unstated duration are considered in default, the data actually show a small decrease in the second year.

identical in terms of their observables, Y and L. Heuristically, a decrease in $\underline{s}$ keeps the position of the households constant, but yields a leftward shift in the graphed decision. A small decrease in $\underline{s}$ can yield qualitatively different responses to the five households labeled. Household I's income is lower than expected, and so would respond to small decrease in $\underline{s}$ by borrowing to the limit and increasing consumption. Household II is a household that had higher than expected income. Without the intervention, the household invests and is not constrained in its consumption. Given the lower $\underline{s}$, it does not borrow, but nevertheless increases its consumption. Given the lower borrowing constraint in the future, it no longer requires as large a bufferstock today. Household III, though not investing, will similarly increase consumption without borrowing by reducing its bufferstock given a small decrease in $\underline{s}$. Thus, in terms of consumption, Household I-III would increase consumption, and Households II and III would do so without borrowing. If these households were the only households, the model would deliver the suprising result that consumption increases more than credit, but Households IV and V work against this. Household IV is a household in default. A small decrease in $\underline{s}$ would have no affect on its consumption or investment, but simply increase the indebtedness of the household and reduce the amount of credit that would have been defaulted. Finally, Household V is perhaps the target household of microcredit rhetoric because given a small increase in credit. But if (as drawn) the household will invest in a sizable project, it will finance this by not only increasing its borrowing but also by reducing its current consumption. One can also see that the effects of changes in $\underline{s}$ are not only heterogeneous, but also nonlinear. For example, if the decrease in $\underline{s}$ were large enough relative to i^*, Household V would not only invest but also increase consumption.

Quantitatively, draws from the distributions of U and i^* determine the scattering of households in each village across Figure 3. The high level of transitory income growth volatility lead to a high variance in U, hence a diffuse distribution in the P dimension. There are a substantial number of defaulters in the baseline data, but the changes in $\underline{s}$ lead to fewer defaulters (like Household IV) and more hand-to-mouth consumers (like Household I). Similarly,the low investment probability but sizable average investment levels in the data lead to high mean and variance of the i^* distribution. Most households have very large projects, but investment is relatively infrequent. Most actual investments are small, but there is great dispersion.[30] Hence, while some households lie close enough to the threshold that changes in $\underline{s}$ induce investment, the vast majority of these investments are small. Investment levels overall are driven by the few very large i^* investments (e.g., a large truck or a warehouse) for which the density of households lying just left of the threshold is relatively small. Thus, few households resemble Household V.

Since a lower $\underline{s}$ can never reduce investment, the theoretical effect of increased liquidity on investment is clear. It is simply that the samples are too small to measure it. Given enough households, a small amounts of credit available will eventually decide whether a very large investment is made or not, and this will occurs more often the larger the decrease in $\underline{s}$. Indeed, when the 500 samples are pooled together, the pooled estimates of 0.40 (standard error=0.04) for the first year is highly significant. The estimate is also sizable. Given the average credit injection per household, this would be an increase in investment of 3800 baht per household (relative to a pre-sample average of 4600 baht/household).

5.3 Normative Analysis

We evaluate the benefits of the Million Baht program by comparing its benefits to a simple liquidity transfer. As our analysis of Figure 3 indicates reductions in $\underline{s}$ (leftward shifts in the policy function from the Million Baht program) are similar to increases in liquidity (rightward shifts in the households from the transfer). Both provide additional liquidity.

The advantage of the Million Baht program is that it provides more than a million baht in potential liquidity ($\Delta \underline{s} P$). That is, (by construction) borrowers increase their credit by roughly a million baht, but non-borrowers also benefit from the increased potential liquidity from the relaxed borrowing constraint. Those that borrow have access to a disproportionate amount of liquidity then they would if the money were distributed equally as transfers.

[30] Another way of interpreting this is that most households do not have potential projects that are of the relevant scale for microfinance. Households with unrealistically large projects may correspond, in the real world, to households that simply have no potential project in which to invest.

The disadvantage of the Million Baht program is that it provides this liquidity as credit, and hence there are interest costs which are substantial given $r = 0.054$. A household that receives a transfer of, say, 10,000 baht earns interest on that transfer relative to a household that has access to 10,000 baht in credit, even if it can be borrowed indefinitely.

The relative importance of these two differences depends on household's need for liquidity. Consider again the household in Figure 3. Household II and III, who are not locally constrained, benefit little from a marginal decrease in $\underline{s}$, since they have no need for it in the current period, and may not need it for quite some time. Households IV, who is defaulting, is actually hurt by a marginal reduction in $\underline{s}$, since the household will now hold more debt, and be forced to pay more interest next period. On the other hand, Households I and V benefit greatly from the reduction in $\underline{s}$, since both are locally constrained.

A quantitative cost benefit analysis is done by comparing the cost of the program (the reduction in $\underline{s}$) to a transfer program (an increase in l) that is equivalent in terms of providing the same expected level of utility (given $L_{n,t}$ and $Y_{n,t}$ in 2001, just before the program is introduced). That is, we solve the equivalent transfer T_n for each household using the following equation:

$$E\left[V(L, P, I^*; \underline{s}_v^{mb})|L_{n,5,v}, Y_{n,5,v}\right] = E\left[V(L + T_n, P, I^*; \underline{s})|L_{n,5,v}, Y_{n,5,v}\right]$$

The average equivalent liquidity transfer per household in the sample is just 8200 baht, whereas the direct costs of the Million Baht Program costs 10,100 baht per household in the sample.[31] Thus, a simple transfer could have provided the benefit of the Million Baht program to these households for just 80 percent of the cost. Thus, although the Million Baht program is able to offer the typical household more liquidity (e.g., in the median village, $\Delta\underline{s}P$ =13,400 baht for a household with average income, while the average cost per household in that village is 9100 baht), this benefit is swamped by the interest costs to households.

[31] This includes only the seed fund, and omits any administrative or monitoring costs of the village banks.

5.4 Alternative Structural Analyses

The structural model allows for several alternative analyses including comparison with reduced form predictions, robustness checks with respect to the return on investment R, estimation using post-intervention data, long run predictions and policy counterfactuals. We briefly summarize the results here, but details are available upon request.

5.4.1 Return on Investment

Our baseline value of R was 0.11. Recall that two alternative calibrations of the return on assets were calculated based on the whether our measure of productive assets included uncultivated or community use land ($R = 0.08$) or the value the plot of land containing the home ($R = 0.04$). We redid both the estimation and simulation using these alternative values. For $R = 0.08$, the estimates were quite similar; only a higher β (0.94), a lower r (0.32); and a lower risk aversion (1.12) were statistically different than the baseline. The modeled had even more difficulty matching income growth and volatility, so that the overall fit was substantially worse (J-statistic=200 vs. 113 in the baseline). The simulation regression estimates were nearly identical. For the low value of $R = 0.04$, the estimation required that the return on liquidity be substantially lower than in the data ($r = 0.018$), and that β be substantially higher (0.97) than typical for bufferstock models. The fit was also substantially worse (J-statistic=324). Finally, the regression estimates on the simulated data were qualitatively similar but smaller (e.g., a consumption coefficient of 0.68 in the first year.) Indeed, only the reduction of default in the first year was statistically significant at a 0.05 percent level.

5.4.2 Estimation Using Ex Post Data

In this analysis, rather than use the post-intervention data to test the model using calibrated borrowing constraints, we use it to estimate the new borrowing constraints and better identify the other parameters in the model. We proceed as follows. First, we create a new panel of data by rerunning the data filtering regressions of 4.1.1 using the full sample of data. Second, we specify a reasonably flexible but parametric function for $\underline{s}_{mb}$ in the post-program years:

$$\underline{s}_{mb,v} = \underline{s}_1 + \underline{s}_2 \left(\frac{1}{\text{\# HHs in village}_v} \right)^{\underline{s}_3}$$

where $\underline{s}_1$, $\underline{s}_2$, and $\underline{s}_3$ are the parameters of interest.[32] Third, for the post-program years, we add additional year-specific moments for income growth and income growth volatility; consumption, investment probability, investment, and their interactions with measured income and liquidity ratios; and default. In total, the estimation now includes 39 moments and 14 parameters.

The estimated results from the full sample are strikingly similar to the baseline estimates from the presample and the calibration from the post-sample.[33] The resulting estimates are $\hat{\underline{s}}_1 = -0.06$, $\hat{\underline{s}}_2 = -127.2$, and $\hat{\underline{s}}_3 = -1.49$. The model fit performs well and fails along the similar dimensions as the baseline. Finally, the average, standard deviation, minimum and maximum of $\underline{s}_{mb,v}$ implied by the estimates are -0.32 (-0.28 in baseline calibration), –0.16 (-0.14), -0.94 (-0.91), and -0.07 (-0.09) respectively. The correlation between the two approaches one by construction, since both increase monotonically with village size.The fact that the estimates and calibrated values are nearly identical indicates that on average the simulated predictions of the model approximate a best fit to the actual data.

5.4.3 Long Run Predictions

The differences between $\hat{\alpha}$ estimates in the first and second year of the program indicate that impacts are time-varying. The structural model allows for simulation and longer run horizon estimates of impact. We therefore simulate datasets that include five additional years of data and ran the analogous regressions. Seven years out, none of the $\hat{\alpha}$ estimates

[32] If all households borrowed every period and had identical permanent income, then the extra borrowing per household ($950,000/\#$ HHs in village$_v$) would translate into borrowing constraints with $\underline{s}_1 = \underline{s}$ (the pre-intervention borrowing constraint) $\underline{s}_2 = \frac{950{,}000}{P}$ and $\underline{s}_3 = 1$.

[33] For comparison, the point estimates are $\hat{r} = 0.057$, $\hat{G} = 1.04$, $\hat{\sigma}_N = 0.30$, $\hat{\sigma}_u = 0.34$, $\hat{\sigma}_E = 0.23$, $\hat{\underline{c}} = 0.45$, $\hat{\beta} = 0.923$, $\hat{\rho} = 1.24$, $\hat{\mu}_i = 0.83$, $\hat{\sigma}_i = 5.70$, and $\hat{\underline{s}} = -0.16$.

are significant on average.While the average point estimates are quite small for investment probability (0.23), investment (0.10), and default probability (0.01) relative to the first year, the average $\hat{\alpha}$ for consumption remains substantial (0.58) and close to the estimate in the second year (0.73). In the model, the impacts on consumption fall somewhat after the first year, but there remains a substantial persistent effect. Still, regression estimates that simply measure common α coefficient for all post-program years do not capture any significant impact on consumption in the long run data. This shows the importance of considering the potential time-varying nature of impacts in evaluation.

5.4.4 Policy Counterfactual

From the perspective of policymakers, the Million Baht Village Fund Program may appear problematic along two fronts. Its most discernible impacts are on consumption rather than investment, and it appears less cost-effective than a simple transfer mainly because funds may simply go to prevent default and the increased borrowing limit actually hurts defaulting households. An alternative policy that one might attempt to implement would be to only allow borrowing for investment. We would assume that the village can observe investment, but since money is fungible, it would be unclear whether these investments would have been undertaken even without the loans, in which case the loans are really consumption loans. Since defaulting households cannot undertake investments, it would prevent households in default from borrowing. Nevertheless, such a policy would also eliminate households like Household 1 in Figure 3 from borrowing.

The ability to model policy counterfactuals is another strength of a structural model. In a model with this particular policy, households face the constraint $\underline{s}_v^{mb,alternative}$ in any period in which they decide to invest, while facing the baseline $\underline{s}$ if they decide not to invest. The default threshold is also moved to $\underline{s}_v^{mb,alternative}$, however, to prevent households from investing and borrowing in one period, and then purposely not investing in the next period in order to default. Under this policy, the new borrowing constraints are even lower (averaging -0.67 vs. -0.28 in the actual policy) but only for those who borrow. The new range of borrowing constraints is from -0.16 to -4.78.

The policy increases both the impact on consumption and increase the impact on investment. Pooling all 500 simulated samples yields a significant estimate for consumption that is similar to the actual million baht intervention (1.40 vs. 1.38 in the first year). It also yields a much larger and significant estimate for investment levels (0.62 in the first year). Clearly, the counterfactual policy channels funds only to investors and so it is able to relax borrowing constraints much more substantially for investors, and in turn to help with large investments. but relaxes the borrowing constraint even more for investing households than the actual policy. Finally, the negative impact on default no longer exists. Although this policy offers less flexibility for constrained households who would rather not invest, the benefits are larger to defaulters and investors help outweigh some of this loss. There is much more variation in the benefits across households (e.g., the standard deviation of the equivalent transfer is 14,000 baht in this counterfactual vs. 11,000 in the baseline policy), but the average equivalent transfer is actually lower (7500 vs. 8200).

6 Conclusions

We have developed a model of bufferstock saving and indivisible investment, and used it to evaluate the impacts of the Million Baht program as a quasi-experiment. The correct prediction of consumption increasing more than one for one with the credit injection is a "smoking gun" for the existence of credit constraints, and is strong support for the importance of bufferstock savings behavior. Nevertheless, the microfinance intervention appears to be less cost effective than a simpler transfer program because it saddles households with interest payments. Finally, we have emphasized the relative strengths of a natural experiment, a structural model, and reduced form regressions.

One limitation of the model is that although project size is stochastic, the quality of investments, modeled through R, is constant across projects and households. In the data, R varies substantially across households. Heterogeneity in project quality may be an important dimension for analysis, since microfinance may change the composition of project quality. The process for project sizes was also extremely stylized. Also, potential projects may be arrive less often, be less transient (which allows for important anticipatory savings behavior as in Buera, 2008), and multiple projects may be ordered by their profitability. Such extensions might help explain the gap between positive predicted impacts on investment probability, but no impact in the data.

Related, the analysis has also been purely partial equilibrium analysis of household behavior. For a large scale intervention, one might suspect that general equilibrium effects on income, rates of return to investment, and interest rates on liquidity may be important (see Kaboski and Townsend, 2007). Finally, we did not consider the potential interactions between villagers or between villages, nor was the intermediation mechanism explicitly modeled. These are all avenues for future research.

References

[1] Aaronsen, D., Agarwal, S. and French, E. "The Consumption Response to Minimum Wage Increases." FRB Chicago working paper 2007-23, May, 2008

[2] Ahlin, C. and Townsend, R. "Using Repayment Data to Test Across Models of Joint Liability Lending." mimeo, 2000

[3] Aiyagari, S. R., "Uninsured Idiosyncratic Risk and Aggregate Saving." *Quarterly Journal of Economics,* 109 (August 1994): 659-684

[4] Alvarez, F. and Stokey, N. "Dynamic Programming with Homogeneous Functions", *Journal of Economic Theory*, 82 (1998): 167-189

[5] Aportela, F. "Effect of Financial Access on Savings by Low-Income People." mimeo, Massachusetts Institute of Technology, 1998

[6] Banerjee, A. and Newman, A. "Occupational Choice and the Process of Development." *Journal of Political Economy, (*April 1993*): 274-98*

[7] Banerjee, A. and Duflo, E. "Do Firms Want to Borrow? Testing Credit Constraints Using a Direct Lending Program." mimeo, Massachusetts Institute of Technology, May 2002

[8] Blossfeld, H. and Shavit,Y. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries. (*Boulder, Co: Westview Press), 1993.

[9] Buera, F. J. "Persistency of Poverty, Financial Frictions, and Entrepreneurship." mimeo, Northwestern University, 2006

[10] Carroll, C. D. "Buffer Stock Saving and the Lifecycle Permanent Income Hypothesis." *Quarterly Journal of Economics, 107 (1) 1997*

[11] Carroll, C. D. "Buffer Stock Saving: Some Theory." Manuscript, The Johns Hopkins University, 2004

[12] Chiappori, P., Townsend, R. and Yamada, H. "Sharing Wage Risk.", mimeo, 2008

[13] Carroll, C. D. and Samwick, A. A. "The Nature of Precautionary Wealth." *Journal of Monetary Economics,* 40 (1) 1997

[14] Coleman, B. "Microfinance in Northeast Thailand: Who Benefits and How Much?" Asian Development Bank ERD Working Paper Series No. 9, 2001

[15] Coleman, B. "The Impact of Group Lending in Northeast Thailand." *Journal of Development Economics*, (October 1999): 105-141

[16] Galor, O. and Zeira, J. "Income Distribution and Macroeconomics." *Review of Economic Studies*, 60 (1 1993): 35-52

[17] Gine, X. and R. Townsend "Evaluation of Financial Liberalization: A Model with Constrained Occupation Choice." mimeo, University of Chicago, 2001

[18] Gourinchas, P.-O., and Parker, J."Consumption over the Life Cycle." *Econometrica,* 71 (January 2002)*:*47-91.

[19] Heckman, J. J., Urzua, S., Vytlacil, E., 2004. "Understanding Instrumental Variables in Models with Essential Heterogeneity". Stanford University, Stanford, CA. http://www.stanford.edu/~vytlacil/

[20] Kaboski, J. and Townsend, R. "The Impacts of Credit on Village Economies." mimeo, 2007

[21] Kaboski, J. and Townsend, R. "Borrowing and Lending in Semi-Urban and Rural Thailand." mimeo, University of Chicago, 1998

[22] Kaboski, J. and Townsend, R. "Policies and Impact: An Analysis of Village Microfinance Institutions." mimeo, University of Chicago, 2003

[23] Karlan, D. and Zinman, J. "Expanding Credit Access: Using Randomized Supply Decisions To Estimate the Impacts." mimeo, August, 2006

[24] Lloyd-Ellis, H. and Bernhardt, D. "Enterprise, Inequality, and Economic Development." *Review of Economic Studies*, (January 2000): 147-68

[25] Lochner, L. and Monge-Naranjo, A. "The Nature of Credit Constraints and Human Capital." NBER working paper W13912, April 2008

[26] Owen, A. and Weil, D. "Intergenerational Earnings Mobility, Inequality, and Growth." NBER Working Paper 6070 (June 1997)

[27] Paulson, A. and Townsend, R. M. "Entrepreneurship and Liquidity Constraints in Rural Thailand." mimeo, 2000

[28] Rosenzweig, M. and Wolpin. K "Natural 'Natural Experiments' in Economics." *Journal of Economic Literature* 38 (December 2000): 827-874

[29] Samphantharak, K. *Internal Capital Markets in Business Groups*. Ph. D. thesis, University of Chicago: Chicago, 2003

[30] Samphantharak, K. and Townsend, R. ““Households as Corporate Firms: Constructing Financial Statements from Integrated Household Surveys” mimeo, 2006

[31] Todd, P. and Wolpin, K. “Using Experimental Data to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico.” working paper, University of Pennsylvania, 2003

[32] Townsend, R., principal investigator with Paulson, A., Sakuntasathien, S., Lee, T. J., and Binford, M., Questionnaire Design and Data Collection for NICHD Grant *Risk, Insurance and the Family* and NSF grants, The University of Chicago, 1997

[33] World Bank “KUPEDES: Indonesia's Model Small Credit Program.” OED Precis, Number 104, February 1996, 4 pages

[34] Wright, M. “Investment and Growth with Limited Commitment.” mimeo, October 2002

[35] Zeldes, S. “Consumption and Liquidity Constraints: An Empirical Investigation.” *Journal of Political Economy*, 97 (March 1989): 305-46

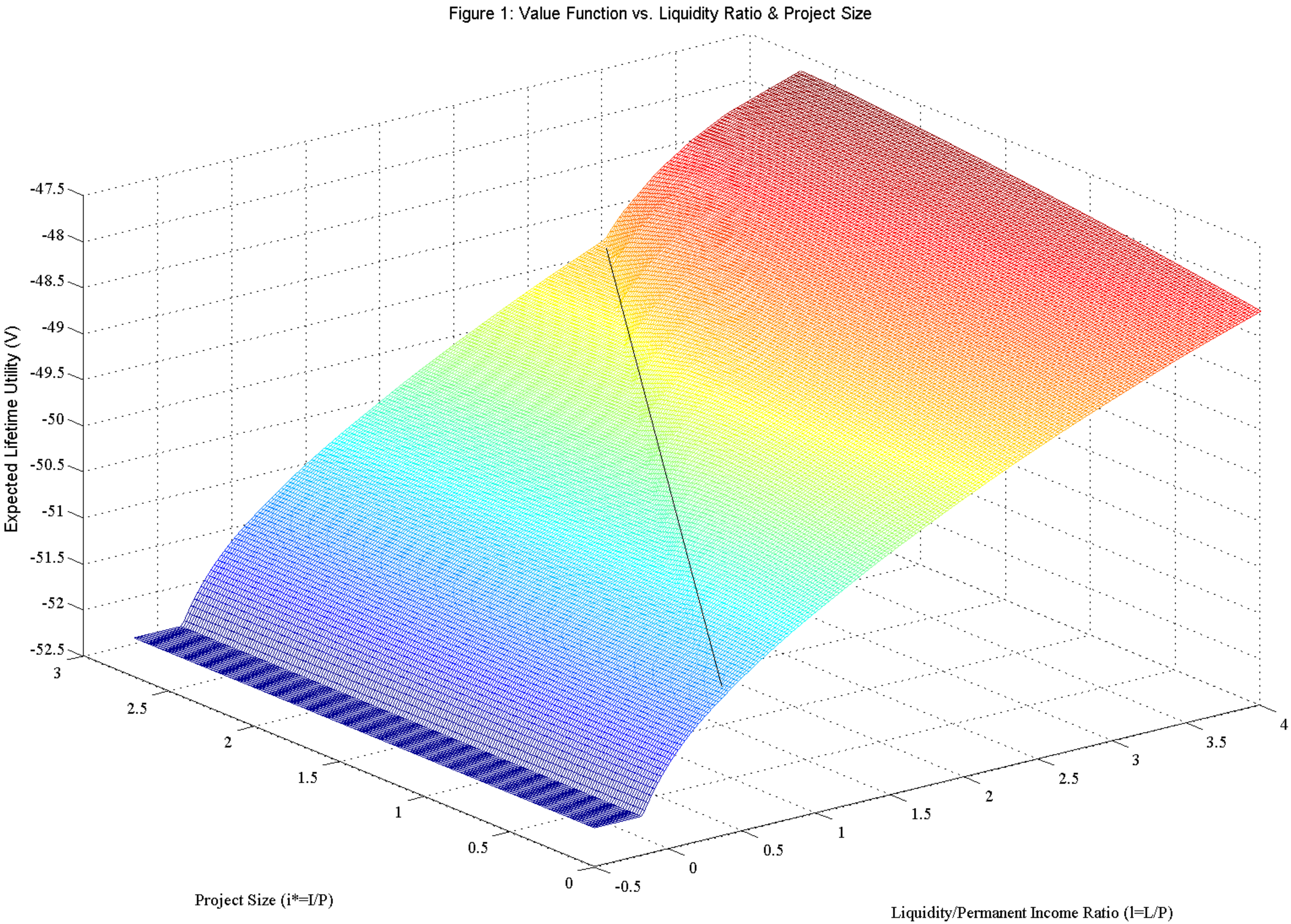

Figure 1: Value Function vs. Liquidity Ratio & Project Size
Expected Lifetime Utility (V)
-47.5
-48
-48.5
-49
-49.5
-50
-50.5
-51
-51.5
-52
-52.5
3
2.5
2
1.5
1
0.5
0
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
Project Size (i*=I/P)
Liquidity/Permanent Income Ratio (l=L/P)

Figure 2: Consumption Policy for Fixed i*, Baseline and Reduced Borrowing Constraint

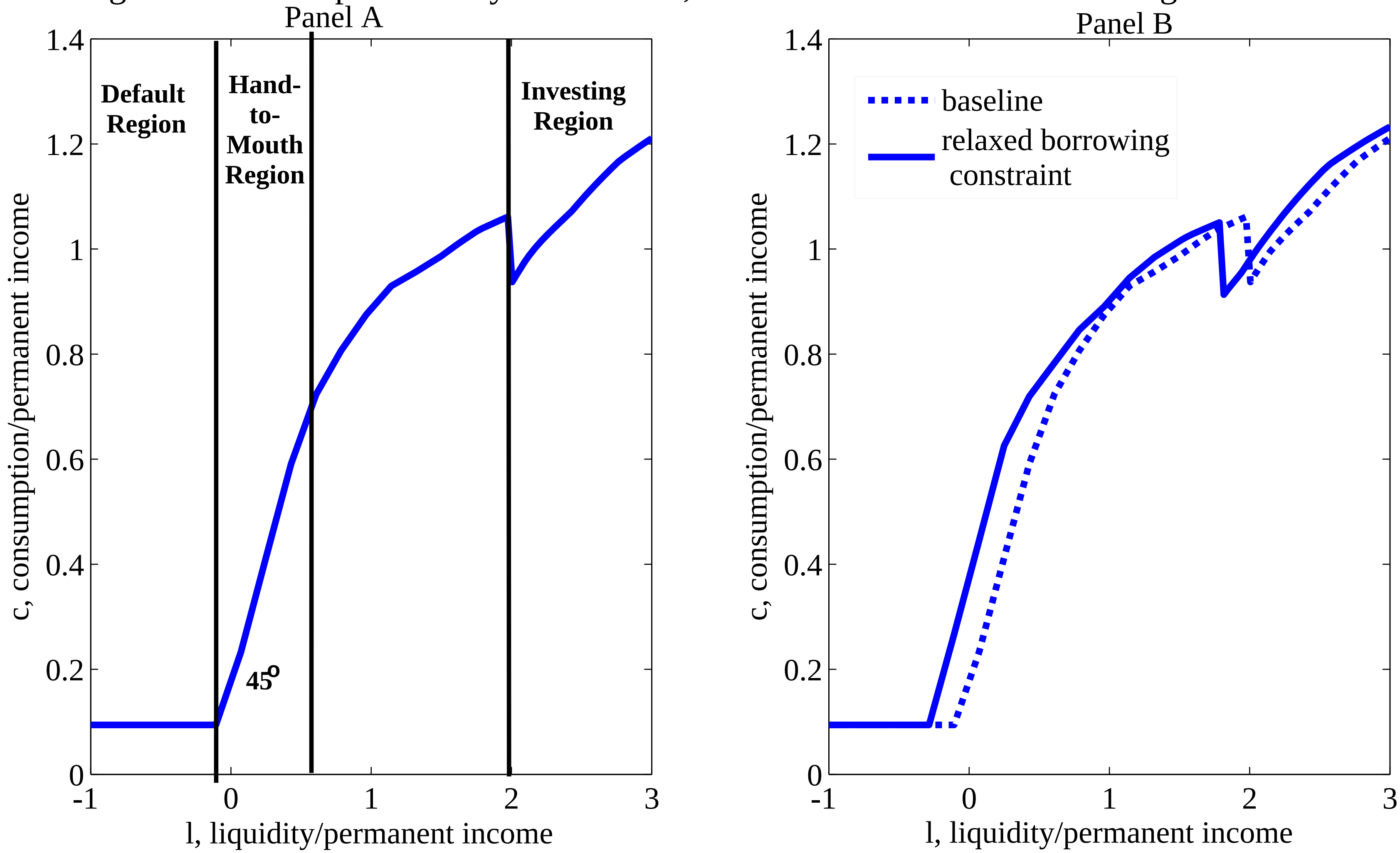

Figure 3: Consumption Policy as a Function of Liquidity and Project Size

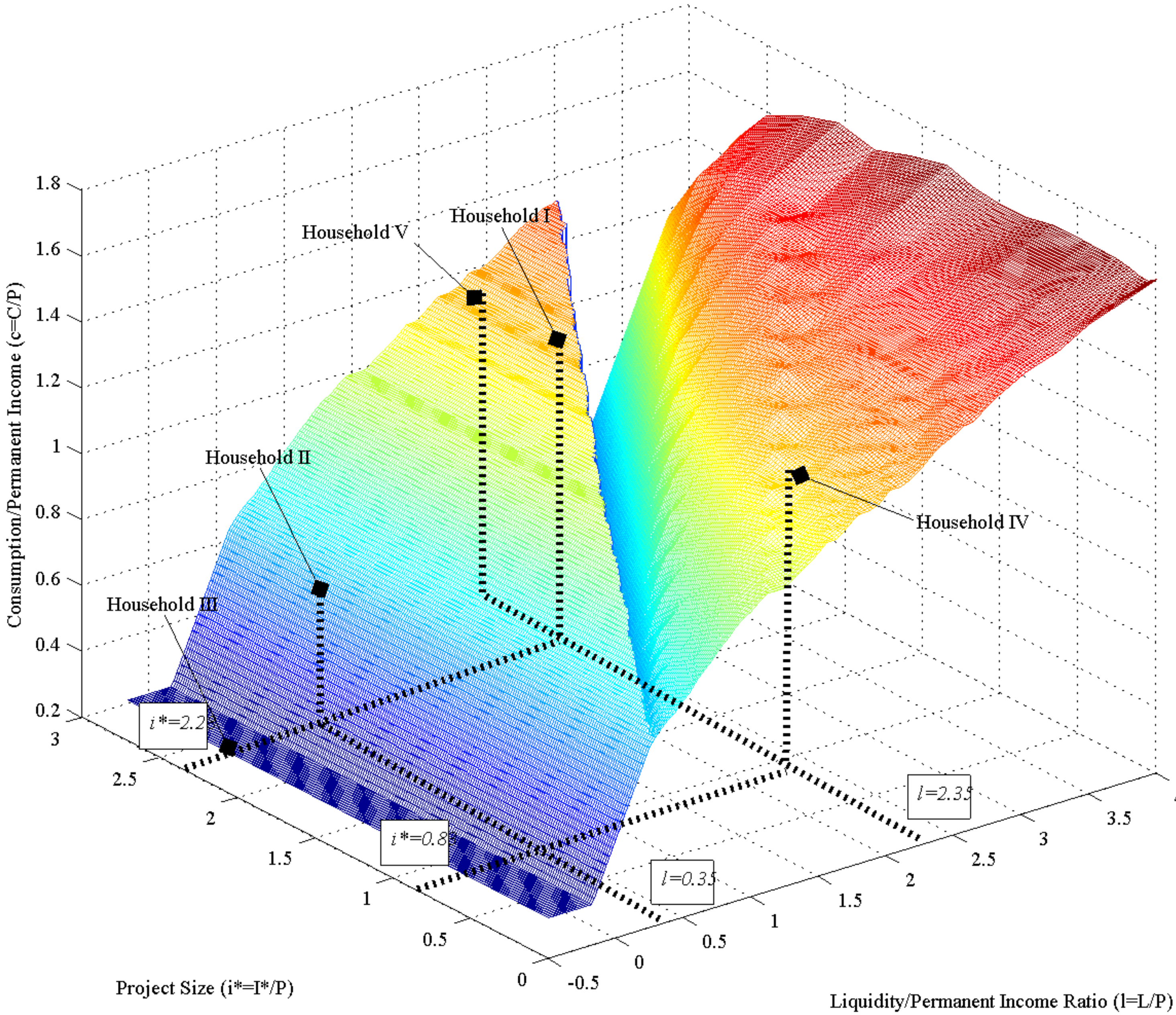

Table 1: Summary Statistics of Pre-Intervention Household Data

Variable	Obs	Mean	Std. Dev.	Min	Median	Max
Primary Variables:						
Non-Interest Household Income*	3575	87200	202000	500	50300	6255500
Log Growth of Income*	2860	0.04	0.98	-4.94	0.01	10.28
Household Consumption*	3575	75200	93000	750	49800	1370300
Dummy Variable for Agr/Business Investment	3575	0.12	0.34	0	0	1
Value of Agr./Business Investment*	3575	4760	30200	0	0	715700
Dummy Variable for Short-Term Default	2860	0.194	0.395	0	0	1
Short-Term Credit*	3575	17900	51100	0	0	1021000
Interest Paid*	3575	1300	3900	0	0	108400
Liquid Savings*	2860	25000	132000	0	5100	4701600
Interest Earned*	3575	700	7200	0	0	18000
Number of Households in Village	3575	166	295	21	110	3194
Regressors for Demographic/Cylical Variation:						
Number of Male Adults	3575	1.46	0.9	0	1	7
Number of Female Adults	3575	1.56	0.75	0	1	6
Number of Children	3575	1.59	1.21	0	1	9
Dummy Variable for Male Head of Household	3575	0.74	0.44	0	1	1
Years of Education of Head of Household	3575	6	3	0	7	15
Age of Head of Household	3575	41	15	22	40	84

* All values are in baht deflated to 1999. The 1999 PPP conversion rate is 31.6 baht/dollar.

Table 3: Identification -
Partial Derivatives of Moments With Respect to Parameters

Moments	r	σ_N	σ_U	σ_E	G	$\underline{c}$	β	ρ	μ_i	σ_i	$\underline{s}$
ε_s	-6.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ε_{cr}	-10.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ε_g	-1.1	1.3	-3.6	-0.5	7.0	1.7	22.0	-2.0	-3.6	0.1	1.4
$\varepsilon_{v,1}$	0.2	-0.7	-0.5	-0.3	-0.8	3.3	0.3	-0.7	1.8	-0.2	0.1
$\varepsilon_{v,2}$	1.5	-1.0	-0.5	0.5	-8.4	2.1	17.5	-0.4	2.4	-0.4	0.8
$\varepsilon_{v,3}$	0.2	-2.1	-0.6	0.5	-3.4	0.8	4.3	0.5	1.5	0.0	-3.9
ε_C	1.4	0.7	0.1	-0.3	-1.6	-0.5	3.3	-7.1	0.0	0.0	0.7
ε_C*lnY	-3.7	-1.8	-0.2	0.7	4.3	1.4	-8.9	18.6	-0.1	0.0	-1.8
ε_C*L/Y	-0.1	-0.1	0.0	0.0	0.1	0.1	-0.4	0.8	0.0	0.0	-0.1
ε_D	52.5	15.4	-0.7	0.0	-40.0	-1.1	-16.3	1.6	-0.5	0.0	-0.6
ε_D*lnY	-158.3	-45.9	2.0	0.1	119.4	3.3	48.8	-4.7	1.4	-0.1	1.8
ε_D*L/Y	-23.6	-9.0	1.7	0.0	18.1	0.6	8.0	-0.7	0.3	0.0	0.3
ε_I	28.7	8.2	0.1	-0.1	-21.8	-0.5	-6.2	1.0	-0.4	0.0	-0.3
ε_I*lnY	-82.2	-23.3	-0.2	0.3	62.3	1.3	17.7	-2.7	1.0	-0.1	0.9
ε_I*L/Y	-10.1	-2.3	0.0	0.4	7.4	0.2	3.4	-1.7	0.0	0.0	0.1
ε_{DEF}	0.0	0.0	-0.9	0.0	0.0	-3.6	0.0	0.0	0.0	0.0	-3.6

Table 2: Parameter Estimates and Model Fit

Parameter Estimates

Parameter	Estimate	Std. Err.
r	0.054	0.003
σ_N	0.31	0.11
σ_U	0.42	0.07
σ_E	0.15	0.09
G	1.047	0.006
$\underline{c}$	0.52	0.01
β	0.926	0.006
ρ	1.20	0.01
μ_i	1.47	0.09
σ_i	6.26	0.72
$\underline{s}$	-0.08	0.03

Pre-Intervention Averages

Variable	Data	Model
C_t	75,200	75,800
D_t	0.116	0.116
I_t	4600	4600
DEF_t	0.194	0.189
$\ln(Y_{t+1}/Y_t)$	0.044	0.049

Test for Overidentifying Restrictions

	Actual Value	0.05% Value
J-Statistic	113.5	12.6

Table 4: Reduced Form Regression Estimates: Actual Data vs. "Million Baht" Simulated Data

	Consumption		Investment Probability		Investment		Default Probability		Income Growth	
Actual Data	$\gamma_{C,2002}$	$\gamma_{C,2003}$	$\gamma_{D,2002}$	$\gamma_{D,2003}$	$\gamma_{I,2002}$	$\gamma_{I,2003}$	$\gamma_{DEF,2002}$	$\gamma_{DEF,2003}$	$\gamma_{\Delta \ln Y,2002}$	$\gamma_{\Delta \ln Y,2003}$
"Impact" Coefficient*	**1.39**	**0.90**	**6.3e-6**	-0.2e-6	-0.04	-0.17	**-5.0e-6**	**6.4e-6**	-9.4e-6	**12.6e-6**
Standard Error	**0.39**	**0.39**	**2.4e-6**	2.4e-6	0.19	0.19	**2.4e-6**	**2.4e-6**	6.1e-6	**6.1e-6**
Simulated Data										
Average "Impact" Coefficient*	**1.10**	0.73	**5.6e-6**	3.6e-6	0.41	0.35	**-9.0e-6**	-0.2e-6	0.3e-6	0.3e-6
Average Standard Error	**0.48**	0.48	**2.5e-6**	2.5e-6	0.23	0.23	**2.3e-6**	2.3e-6	5.9e-6	5.9e-6
Fraction Rejecting 5% Chow Test**	0.01		0.02		0.28		0.07		0.05	

*The impact coefficient is the coefficient on 1,000,000/number of households in the village interacted with a year dummy, the credit injection per household.

**This is the fraction of simulations where a Chow test rejects at a 5 percent significance level that the coefficients in the actual and simulated data are the same, once actual data and simulated data are pooled.

Bold face represents significance at a 5 percent level.