Dynamic Financial Constraints: Distinguishing Mechanism Design from Exogenously Incomplete Regimes*

Alexander Karaivanov
Simon Fraser University

Robert M. Townsend
MIT

January 4, 2012

Abstract

We formulate and solve a range of dynamic models of constrained credit/insurance that allow for moral hazard and limited commitment (and in some runs include hidden output and unobserved capital and investment). We compare them to full insurance and exogenously incomplete financial regimes (autarky, saving only, and borrowing and lending in a single asset). We develop computational methods based on mechanism design, linear programming, and maximum likelihood to estimate, compare, and statistically test these alternative dynamic models with financial/information constraints. Our methods can use both cross-sectional and panel data and allow for measurement error and unobserved heterogeneity. We estimate the models using data on Thai households running small businesses in two separate samples. We find that for the rural sample, the saving only and borrowing regimes provide the best fit using data on consumption, business assets, investment, and income from rural Thailand. Family and other networks are helpful in consumption smoothing there, as in a moral hazard constrained regime. In contrast, in urban areas, we find financial/information regime that is decidedly less constrained, with the moral hazard model fitting best even in combined business and consumption data. We run numerous robustness check in the Thai data, and in Monte Carlo simulations, and compare our maximum likelihood criterion with results from other metrics.

Keywords: financial constraints, mechanism design, structural estimation and testing

*We thank H. Cole, L. Hansen, A. Hortacsu, P. Lavergne, E. Ligon, T. Magnac, D. Neal, W. Rogerson, S. Schulhofer-Wohl, K. Wolpin, Q. Vuong, and audiences at Berkeley, Chicago, Duke, MIT, Northwestern, Penn, Santa Barbara, Toulouse, Stanford, the SED Conference, the European Meeting of the Econometric Society, and the Vienna Macroeconomics Workshop for many insightful comments and suggestions. Financial support from the NSF, NICHD, the John Templeton Foundation, SSHRC, and the Bill and Melinda Gates Foundation through a grant to the Consortium on Financial Systems and Poverty at the University of Chicago is gratefully acknowledged.

1 Introduction

We compute, estimate, and contrast the consumption and investment behavior of risk averse households running small non-farm and farm businesses under alternative financial and information environments, including exogenously incomplete markets settings (autarky, savings only, non-contingent borrowing) and endogenously constrained settings (moral hazard, limited commitment, unobserved output, unobserved investment), both relative to full insurance. We analyze in what circumstances these financial regimes can be distinguished in consumption and income data, in investment and income data or both. More generally, we propose and apply methods for structural estimation of dynamic mechanism design models. We use the estimates to statistically test the alternative models against each other using actual data on Thai rural and urban households and data simulated from the models themselves. We conduct numerous robustness checks, including using data not part of the estimation to compare predictions at the estimated parameters, and identify features in the data that drive the results.

With few exceptions the existing literature maintains a dichotomy, also embedded in the national accounts: households are consumers and suppliers of market inputs, whereas firms produce and hire labor and other factors. This gives rise, on the one hand, to a large literature which studies household consumption smoothing.[1]

On the other hand, the consumer-firm dichotomy gives rise to an equally large literature on investment.[2] Mostly, firms are modeled as risk neutral maximizers of expected discounted profits or of dividends to owners. There are also works attempting to explain stylized facts on firm growth, with higher mean growth and variance in growth for small firms, e.g. Cooley and Quadrini (2001), among others. The more recent works by Albuquerque and Hopenhayn (2004) and Clementi and Hopenhayn (2006) introduce either private information or limited commitment but maintain risk neutrality.[3] Here we set aside for the moment the issues of heterogeneity in technologies and firm growth and focus on a benchmark with financial constraints, investment and consumption data thinking of households as firms. The literature that is closest to our paper, and complementary with what we are doing, features risk averse households as firms but largely *assumes* that certain markets or contracts are missing.[4] The methods of our paper might indicate how to build upon these papers, possibly with alternative

[1] There is voluminous work estimating the permanent income model, the full risk sharing model, buffer stock models (Zeldes, 1989; Deaton and Laroque, 1996) and, lately, models with private information (Phelan 1994; Ligon, 1998; Werning, 2001) or limited commitment (Ligon, Thomas and Worrall, 2002).

[2] For example, there is the adjustment costs approach of Abel and Blanchard (1983) and Bond and Meghir (1994) among many others. In industrial organization, Hopenhayn (1992) and Ericson and Pakes (1995) model the entry and exit of firms with Cobb-Douglas or CES production technologies where investment augments capital with a lag and output produced from capital, labor and other factors is subject to factor-neutral Markov technology shocks.

[3] Applied general equilibrium models feature both consumption and investment in the same context, as Rossi-Hansberg and Wright (2007), but there the complete markets hypothesis justifies, within the model, a separation of the decisions of households from the decisions of firms. Alem and Townsend (2009) provide an explicit derivation of full risk sharing with equilibrium stochastic discount factors, rationalizing the apparent risk neutrality of households as firms making investment decisions.

[4] For example, Cagetti and De Nardi (2006) follow Aiyagari (1994) in their study of inequality and assume that labor income is stochastic and uninsurable, while Angeletos and Calvet (2007) and Covas (2006) in their work on buffer stock motives and macro savings rates feature uninsured entrepreneurial risk. In the asset pricing vein, Heaton and Lucas (2000) model entrepreneurial investment as a portfolio choice problem, assuming exogenously incomplete markets in the tradition of Geanakoplos and Polemarchakis (1986) or Zame (1993).

assumptions on the financial underpinnings.

Indeed, this literature begs the question of how good an approximation are the various assumptions on the financial markets environment, different across the different papers. That is, what would be a reasonable assumption for the financial regime if that part too were taken to the data? We take this view below to see how far we can get. For example, the adjustment costs investment literature may be picking up constraints implied by financing, not adjustment costs per se. The pecking order investment literature (Myers and Majuf, 1984) simply assumes that internally generated funds are least expensive, followed by debt, and finally equity, discussing wedges and distortions. Berger and Udell (2002) also have a long discussion in this spirit, of small vs. large firm finance. They point out that small firms seem to be informationally opaque yet receive funds from family, friends, angels, or venture capitalists, leaving open the nature of the overall financial regime.[5] The empirical work of Fazzari, Hubbard and Petersen (1988) picks up systematic distortions for small firms, but, again, the nature of the credit market imperfection is not modeled, leading to criticisms of their interpretation of cash flow sensitivity tests (Kaplan and Zingales, 2000).[6]

Relative to most of the literature, the methods developed and used in this paper offer the following advantages. First, we solve and estimate fully dynamic models of incomplete markets – this is computationally challenging but captures the complete, both static and dynamic implications of financial constraints on consumption, investment and production. Second, our empirical method can handle any number of type of financial regimes with different frictions. Third, by using a maximum likelihood approach as opposed to reduced form or estimation methods based on Euler equations we are in principle able to estimate a wider set of structural parameters than, for example, those that appear in investment or consumption Euler equations. More generally, the MLE approach allows us to capture more dimensions of the joint distribution of variables in the data (consumption, income, investment, capital) than those captured by other methods, as opposed to only particular dimensions such as consumption-income comovement or cash flow/investment correlations. Fourth, on the technical side, our use of linear programming methods allows us to deal more easily with non-convexities which are common in the endogenously incomplete markets settings than alternative approaches using first order conditions. The combination between linear programming and MLE allows for an easy and direct mapping between model solutions (already in probabilistic form) and likelihoods which may be unavailable using other solution or estimation methods. Our methods are generally applicable to various dynamic discrete choice decision problems by first writing them as linear programs and then mapping the solutions into likelihoods.

Our methods follow logically from Paulson, Townsend and Karaivanov (2006) where we model, estimate, and test whether moral hazard or limited liability is the predominant financial obstacle causing the observed positive monotonic relationship between initial wealth and subsequent decision to enter into business. Buera and Shin (2007) extend this to endogenous savings decisions in a model

[5]Bitler, Moskowitz and Vissing-Jorgensen (2005) argue likewise that agency considerations play important role.

[6]Under the null of complete markets there should be no significant cash flow variable in investment decisions, but the criticism is that when the null is rejected, one cannot infer the degree of imperfection of financial markets from the magnitude of the cash flow coefficient. One needs to explicitly model the financial regime in order to make an inference, which is what we do in this paper.

with limited borrowing. Here, again, we abstract from occupational choice and focus much more on the dynamics. The recent work of Schmid (2008) is also an effort to estimate a dynamic model of financial constraints using regressions on model data, not maximum likelihood as here. Dubois et al. (2008) estimate semi-parametrically a dynamic model with limited commitment to explain the patterns of income and consumption growth in Pakistani villages, nesting the case of complete markets and the case where only informal agreements are available. Finally, Kinnan (2009) uses non-parametric methods to test inverse Euler equations and other implications of moral hazard, limited commitment, and unobserved-output financial regimes.

We naturally analyze the advantages of using a combination of data on consumption and the smoothing of income shocks with data on the smoothing of investment from cash flow fluctuations, in effect filling the gap in the dichotomy of the literature. In estimating both exogenously incomplete and endogenous information-constrained regimes we also break new ground. The only other similar efforts of which we are aware are Meh and Quadrini (2006), who compare and contrast numerically a bond economy to an economy in which unobserved diversion of capital creates an incentive constraint, and Attanasio and Pavoni (2008) who estimate and compare the extent of consumption smoothing in the permanent income model to that in a moral hazard model with hidden savings (see also Karaivanov, 2008). Krueger and Perri (2010) use data on income, consumption and wealth from Italy (1987-2008) and the PSID (2004-2006) to compare and contrast the permanent income hypothesis (borrowing and lending in a risk-free asset) vs. a model of precautionary savings with borrowing constraints and conclude the former explains the dynamics of their data better.

In this paper we focus on whether, and in what circumstances, it is possible to distinguish financial regimes, depending on the data used. To that end we perform tests in which we have full control, that is, we know what the financial regime really is, using data generated from the model. Our paper is thus both a conceptual and methodological contribution. We show how all the financial regimes can be formulated as linear programming problems, often of large dimension, and how likelihood functions, naturally in the space of probabilities/lotteries, can be estimated. We allow for measurement error, the need to estimate the underlying distribution of unobserved state variables, and the use of data from transitions, before households reach steady state.

We apply our methods to a featured emerging market economy – Thailand – to make the point that what we offer is a feasible, practical approach to real data when the researcher aims to provide insights on the source and nature of financial constraints. We chose Thailand for two main reasons. First, our data source (the Townsend Thai surveys) includes panel data on both consumption and investment and this is rare. We can thus see if the combination of consumption and investment data really helps make a difference. Second, we also learn about potential next steps in modeling financial regimes. We know in particular, from other work with these data, that consumption smoothing is quite good, that is, it is sometimes difficult to reject full insurance, in the sense that the coefficient on idiosyncratic income, if significant, is small (Chiappori, Schulhofer-Wohl, Samphantharak, and Townsend, 2008). We also know that investment is sensitive to income, especially for the poor, but on the other hand this is to some extent overcome by family networks (Samphantharak and Townsend, 2009). Finally, there is a seeming divergence between high rate of return households, who seem constrained in scale, and low rate of return households, who seemingly should be doing something else with their funds. In short, intermediation is imperfect but varies depending on the dimension chosen.

While we keep these data features in mind, we remain neutral in what we expect to find in terms of the best-fitting theoretical model. Hence, we test the full range of regimes from autarky to full information against the data. We are interested in how these same data look when viewed jointly though the lens of each of the various financial regimes modeled here. We also want to be assured that our methods which use grid approximations, measurement error, estimation of unobserved distribution of utility promises, and transition dynamics are, as a practical matter, applicable to actual data. This is our primary intent, to create an operational methodology for estimating and comparing across different dynamic models of financial regimes that can be taken to data from various sources. We focus on the Thai application, first, but use Monte Carlo simulation methods and a variety of robustness checks, including with data or metrics not used in the estimation, to make sure our methods are working properly.

We find that by and large out methods work with the Thai data. In terms of the regime that fits the Thai rural data best, we echo previous work which finds that investment is not smooth, can be sensitive to cash flow fluctuations and the capital stock is persistent. Indeed, we find that investment and income data alone are most consistent with the borrowing and lending or savings only regimes, with ties depending on the specification and this is true as well with the combined consumption and investment data. We also echo previous work which finds that full risk sharing is rejected, but not by much, and indeed find that the moral hazard regime is not inconsistent with the income and consumption data alone, though often statistically tied with borrowing and lending or savings only, depending on the specification. We find some evidence that family networks move households more decisively toward less constrained regimes.

In the urban data, we find the financial/information regime is less constraining overall. There is still some persistence in the capital stock, though less so than in the rural data, so with production data alone the savings regime again fits best. But the consumption data is even more smooth against income than in the rural data and the moral hazard and even full insurance regimes fit well, with less ties with the more limited saving and borrowing regimes. Overall in the urban data, with combined production and consumption data, the moral hazard regime consistently provides the best fit.

We are also keen to distinguish among the financial/information regimes themselves, and not auxiliary assumptions. So in a major robustness check we establish that imposed parametric production function, with estimated parameters, does not drive our conclusion. The primary specification uses unstructured histograms for input/output data, as our computational methods allow arbitrary functional forms.

We also perform a range of additional runs that confirm the robustness of our results in the Thai data — imposing risk neutrality, imposing no measurement error in the generated data, allowing for quadratic adjustment costs in investment, allowing for limited commitment in the moral hazard and full information regimes, different grid sizes, running on data cleaned from household or year fixed effects, and alternative assets definitions.

In another major robustness check on what we do with actual data, we perform Monte Carlo estimations with model-generated data. In these runs we know what the financial regime really is, and what the true parameter values really are. We find that our ability to distinguish between the financial regimes naturally depends on both the type of data used and the amount of measurement error. With low measurement error we are able to distinguish between almost all regime pairs. As expected, however, higher level of measurement error in the data reduces the power of our model comparison test – some counterfactual regimes cannot be distinguished from the data-generating baseline as well as from each other. For example, using investment/income data, we cannot distinguish between the moral hazard and full information regimes, even though moral hazard generated the data, when there is high measurement error. Using joint data on consumption, investment, and income however does markedly improve the ability to distinguish across the regimes, even with high measurement error in the simulated data. We also incorporate intertemporal data from the model through a panel which also significantly improves the ability to distinguish the regimes, relative to when using single cross-sections. Finally, the simulated data results are shown to be robust to various modifications of the baseline runs — no measurement error, different grid sizes, allowing for heterogeneity in productivity and risk aversion (then ignoring this in estimation so that the model is mis-specified).

Finally, we come back to key features of the data and how the models do in fitting these features and in predicting variables, moments, and time paths not used in the estimation. We display the persistence of capital in the data and in the best fitting financial/information regimes. This makes it clear why the savings-only regime does best when there is substantial persistence (lack of adjustment) as in the rural, and to a less extent urban, data. We also display the rate of return on assets as a function of assets. Again the more limited financial regimes do best in being consistent with the negative relationship in the rural data, i.e., low-asset households have relatively high rates of return and households with higher assets, low rates. The urban data shares some features with data simulated from a less constrained regime (moral hazard). We also simulate the time paths of the best financial regime at the estimated parameters. Then means of consumption, business assets, and income fit quite well. The standard deviations of these variables also fit, though there is more heterogeneity in the actual than model-generated data (removing extreme outliers helps). We also simulate the path for savings in the model and compare, favorably, to 'financial net worth' monthly data not previously utilized in the estimation.

Though the likelihood routines do not match moments in the data, a mean squared error metric based on selected moments picks out the savings regime as best fitting in the rural data, consistent with out MLE results. The urban data show substantially more smoothing and the ad hoc moments criterion picks out a less constrained regime. We also go beyond our MLE and Vuong test approach in running some standard Euler equation GMM tests with the same data. The Bond and Meghir (1994) investment sensitivity to cash flow test rejects the null of no financial constraints. The Ligon (1998) GMM test with consumption data finds the savings / borrowing regimes fit best in the rural data and evidence in favor of moral hazard in the urban data.

2 Theory

2.1 Environment

Consider an economy of infinitely-lived agents heterogeneous in their initial endowments (assets), k_0 of a single good that can be used for both consumption and investment. Agents are risk averse and have time-separable preferences defined over consumption, c, and labor effort, z represented by $U(c, z)$ where $U_1 > 0$, $U_2 < 0$. They discount future utility using discount factor $\beta \in (0, 1)$. We assume that c and z belong to the finite discrete sets (grids) C and Z respectively. The agents have access to a stochastic output technology, $P(q|z, k) : Q \times Z \times K \to [0, 1]$ which gives the probability of obtaining output/income, q from effort level, z and capital level, k.[7] The sets Q and K are also finite and discrete – this could be a technological or computational assumption. Capital, k depreciates at rate $\delta \in (0, 1)$. Depending on the intended application, the lowest capital level ($k = 0$) could be interpreted as a 'worker' occupation (similar to Paulson et al., 2006) or as 'firm exit' but we do not impose a particular interpretation in this paper.

Agents can contract with a financial intermediary and enter into saving, debt, or insurance arrangements. We characterize the optimal dynamic financial contracts between the agents and the intermediary in different financial markets 'regimes' distinguished by alternative assumptions regarding information, enforcement/commitment and credit access. In all financial regimes we study with the exception of the 'hidden output' regime in the Appendix, output q is assumed to be observable and verifiable. However, one or both of the inputs, k and z may be unobservable to third parties, resulting in moral hazard and/or adverse selection problems.

The financial intermediary is risk neutral and has access to an outside credit market with exogenously given and constant over time opportunity cost of funds R. The intermediary is assumed to be able to fully commit to the ex-ante (constrained-) optimal contract with agent(s) at any initial state while we consider the possibility of limited commitment by the agents.

Using the linear programming approach of Prescott and Townsend (1984) and Phelan and Townsend (1991), we model financial contracts as probability distributions (lotteries) over assigned or implemented allocations of consumption, output, effort, and investment (see below for details). There are two possible interpretations. First, one can think of the intermediary as a principal contracting with a single agent/firm at a time, in which case financial contracts specify mixed strategies over allocations. Alternatively, one can think of the principal contracting with a continuum of agents, so that the optimal contract specifies the fraction of agents of given type or at given state that receive a particular deterministic allocation. It is further assumed that there are no aggregate shocks, there are no technological links between the agents, and they cannot collude.

[7] We can easily incorporate heterogeneity in productivity/ability across agents by adding a scaling factor in the production function, as we do in a robustness run. Note also that q, as defined, can be interpreted as income net of payments for any hired inputs other than z and k.

2.2 Financial and information regimes

We write down the dynamic linear programming problems determining the (constrained) optimal contract in many alternative financial regimes which can be classified into two groups. The first group are regimes with exogenously incomplete markets: *autarky (A)*, *saving only (S)*, and *borrowing and lending (B)*. To save space we often use the abbreviated names supplied in the brackets. In these regimes the feasible financial contracts take a specific, exogenously given form (no access to financial markets, a deposit/storage contract, or a non-contingent debt contract, respectively).

In the second group of financial regimes we study, optimal contracts are endogenously determined as solutions to dynamic mechanism-design problems subject to information and incentive constraints. In the main part of this paper we look at two such endogenously incomplete markets regimes – *moral hazard (MH),* in which agents' effort is unobserved but capital and investment are observed, and *limited commitment (LC)* in which there are no information frictions but agents can renege on the contract after observing the output realization. In robustness checks we also introduce and test two additional financial regimes with endogenously incomplete markets (see Appendices A and B for derivations) – *moral hazard with unobserved investment (UI)*, in which a dynamic adverse selection friction is introduced in addition to moral hazard and *hidden output (HO)*, in which output q is unobservable to the intermediary.[8]

All incomplete-markets regimes are compared to what we call the *full information (FI)* benchmark (the 'complete markets' or 'first best' regime). In Section 5.3 we also consider versions of all regimes in which capital changes are subject to quadratic adjustment costs.

2.2.1 Exogenously incomplete markets

Autarky

We say agents are in 'autarky' if they have no access to financial intermediation or storage. They can however choose how much output to invest in production vs. how much to consume. The timeline is as follows. The agent starts the current period with capital $k \in K$ which he invests into production. The initial capital can be also thought of as the agent's beginning-of-period 'wealth'. At this time he also supplies his effort $z \in Z$. At the end of the period output $q \in Q$ is realized, the agent decides on the next period capital level $k' \in K$ (we allow arbitrary downward or upward capital adjustments), and consumes $c \equiv (1-\delta)k + q - k'$. Clearly, k is the single state variable in the recursive formulation of the agent's problem which is relatively simple and can be solved by standard dynamic programming techniques. However, to be consistent with the solution method that we use for the mechanism-design financial regimes where non-linear techniques may be inapplicable due to non-convexities introduced by the incentive and truth-telling constraints (more on this below), we formulate the agent's problem in autarky (and all others) as a dynamic linear programming problem in the joint probabilities (lotteries)

[8] The proofs that the optimal contracting problems can be written in recursive form and that the revelation principle applies follow or can be adapted from Phelan and Townsend (1991) and Doepke and Townsend (2006) and hence are omitted.

over all possible choice variables allocations (q, z, k') given state k,

$$v(k) = \max_{\pi(q,z,k'|k)} \sum_{Q\times Z\times K} \pi(q, z, k'|k)[U((1-\delta)k + q - k', z) + \beta v(k')] \tag{1}$$

The maximization of the agent's value function, $v(k)$ in (1) is subject to a set of constraints on the choice variables, π.[9] First, $\forall k \in K$ the joint probabilities $\pi(q, z, k'|k)$ have to be consistent with the technologically-determined probability distribution of output, $P(q|z, k)$:

$$\sum_{K} \pi(\bar{q}, \bar{z}, k'|k) = P(\bar{q}|\bar{z}, k) \sum_{Q\times K} \pi(q, \bar{z}, k'|k) \text{ for all } (\bar{q}, \bar{z}) \in Q \times Z \tag{2}$$

Second, given that the $\pi(.)$'s are probabilities, they must satisfy $\pi(q, z, k'|k) \geq 0$ (non-negativity) $\forall (q, z, k') \in Q \times Z \times K$, and 'adding-up',

$$\sum_{Q\times Z\times K} \pi(q, z, k'|k) = 1 \tag{3}$$

The policy variables $\pi(q, z, k'|k)$ that solve the above maximization problem determine the agent's optimal effort and output-contingent investment in autarky for each k.

Saving only / Borrowing

In this setting we assume that the agent is able to either only save, i.e., accumulate and run down a buffer stock, in what we call the *saving only (S)* regime; or borrow and save through a competitive financial intermediary – which we call the *borrowing (B)* regime. The agent thus can save or borrow in a risk-free asset to smooth his consumption or investment in Bewley-Aiyagari manner, in addition to what he could do via production and capital alone under autarky. Specifically, if the agent borrows (saves) an amount b, then next period he has to repay (collect) Rb, independent of the state of the world. Involuntary default is ruled out by assuming that the principal refuses to lend to a borrower who is at risk of not being able to repay in any state.[10] By shutting down all contingencies in debt contracts we aim for better differentiation from the mechanism design regimes.

Debt/savings b is assumed to take values on the finite grid B. By convention, a negative value of b represents savings, i.e., in the S regime the upper bound of the grid B is zero, while in the B regime the upper bound is positive. The lower bound of the grid for b in both cases is a finite negative number. The autarky regime can be subsumed by setting $B = \{0\}$. This financial regime is essentially a version of the standard Bewley model with borrowing constraints defined by the grid B and an endogenous income process defined by the production function $P(q|k, z)$.

The timeline is as follows: the agent starts the current period with capital k and savings/debt b and uses his capital in production together with effort z. At the end of the period, output q is realized, the agent repays/receives Rb, and borrows or saves $b' \in B$. He also puts aside (invests in) next period's

[9] In (1) and later on in the paper K under the summation sign refers to summing over k' and not k and similarly for the set W below.

[10] Computationally, this is achieved by assigning high utility penalty in such states.

capital, k' and consumes $c \equiv (1-\delta)k + q + b' - Rb - k'$. The two 'assets' k and b are assumed freely convertible into one another.

The problem of an agent with current capital stock k and debt/savings level b in the S or B regime can be written recursively as:

$$v(k,b) = \max_{\pi(q,z,k',b'|k,b)} \sum_{Q\times Z\times K\times B} \pi(q,z,k',b'|k,b)[U((1-\delta)k + q + b' - Rb - k', z) + \beta v(k',b')] \quad (4)$$

subject to the technological consistency and adding-up constraints analogous to (2) and (3), and subject to $\pi(q,z,k',b'|k,b) \geq 0$ for all $(q,z,k',b') \in Q \times Z \times K \times B$.

2.2.2 Mechanism Design Regimes

Full information

With full information (FI) the principal fully observes and can contract upon agent's effort and investment – there are no private information or other frictions. We write the corresponding dynamic principal-agent problem as an extension of Phelan and Townsend (1991) with capital accumulation. As is standard in such settings (e.g., see Spear and Srivastava, 1987), to obtain a recursive formulation we use an additional state variable – *promised utility*, w which belongs to the discrete set, W. The optimal full-information contract for an agent with current promised utility w and capital k consists of the effort and capital levels $z, k' \in Z \times K$, next period's promised utility $w' \in W$, and transfers τ belonging to the discrete set T. A positive value of τ denotes a transfer from the principal to the agent. The timing of events is the same as in Section 2.2.1, with the addition that transfers occur after output is observed.

Following Phelan and Townsend (1991), the set of promised utilities W has a lower bound, $w_{\min}$ which corresponds to assigning forever the lowest possible consumption, $c_{\min}$ (obtained from the lowest $\tau \in T$ and the highest $k' \in K$) and the highest possible effort, $z_{\max} \in Z$. The set's upper bound, $w_{\max}$ in turn corresponds to promising the highest possible consumption, $c_{\max}$ and the lowest possible effort forever:

$$w^{FI}_{\min} = \tfrac{U(c_{\min}, z_{\max})}{1-\beta} \text{ and } w^{FI}_{\max} = \tfrac{U(c_{\max}, z_{\min})}{1-\beta} \quad (5)$$

The principal's value function, $V(w,k)$ when contracting with an agent at state (w,k) maximizes expected output net of transfers plus the discounted value of future outputs less transfers. We write the mechanism design problem solved by the optimal contract as a linear program in the joint probabilities, $\pi(\tau,q,z,k',w'|w,k)$ over all possible allocations (τ,q,z,k',w'):

$$V(w,k) = \max_{\{\pi(\tau,q,z,k',w'|w,k)\}} \sum_{T\times Q\times Z\times K\times W} \pi(\tau,q,z,k',w'|w,k)[q - \tau + (1/R)V(w',k')] \quad (6)$$

The maximization in (6) is subject to the 'technological consistency' and 'adding-up' constraints:

$$\sum_{T\times K\times W} \pi(\tau, \bar{q}, \bar{z}, k', w'|w, k) = P(\bar{q}|\bar{z}, k) \sum_{T\times Q\times K\times W} \pi(\tau, q, \bar{z}, k', w'|w, k) \text{ for all } (\bar{q}, \bar{z}) \in Q \times Z, \tag{7}$$

$$\sum_{T\times Q\times Z\times K\times W} \pi(\tau, q, z, k', w'|w, k) = 1, \tag{8}$$

as well as non-negativity: $\pi(\tau, q, z, k', w'|w, k) \geq 0$ for all $(\tau, q, z, k', w') \in T \times Q \times Z \times K \times W$.

The optimal FI contract must also satisfy an additional, *promise keeping* constraint which reflects the principal's commitment ability and ensures that the agent's present-value expected utility equals his promised utility, w:

$$\sum_{T\times Q\times Z\times K\times W} \pi(\tau, q, z, k', w'|w, k)[U(\tau + (1-\delta)k - k', z) + \beta w'] = w \tag{9}$$

By varying the initial promise w we can trace the whole Pareto frontier for the principal and the agent. The optimal FI contract is the probabilities $\pi^*(\tau, q, z, k', w'|w, k)$ that maximize (6) subject to (7), (8) and (9).

The full information contract implies full insurance, so consumption is smoothed completely against output, q (conditioned on effort z if utility is non-separable). It also implies that expected marginal products of capital ought to be equated to the outside interest rate implicit in R, adjusting for disutility of labor effort which the planner would take that into account in determining how much capital k to assign to a project. The intermediary/bank (planner) has access to outside borrowing and lending at the rate R, but internally, within its set of customers it can in effect have them 'borrow' and 'save' among each other, i.e., take some output away from one agent who might otherwise have put money into his project and give that to another agent with high marginal product. A lot of this nets out so only the residual is financed with (or lent to) the outside market. In contrast, the B/S regime shuts down such within-group transfers and trades and instead each agent is dealing with the market directly.

Moral hazard

In the moral hazard (MH) regime the principal can still observe and control the agent's capital and investment (k and k'), but he can no longer observe or verify the agent's effort, z. This results in a moral hazard problem. The state k here can be interpreted as endogenous collateral. The timing is the same as in the FI regime. However, the optimal MH contract $\pi(\tau, q, z, k', w'|w, k)$ must satisfy an incentive-compatibility constraint (ICC), in addition to (7)-(9).[11] The ICC states that, given the agent's state (w, k) and recommended effort level $\bar{z}$, capital k', and transfer τ, the agent must not be able to achieve higher expected utility by deviating to any alternative effort level $\hat{z}$. That is, $\forall(\bar{z}, \hat{z}) \in Z \times Z$ we must

[11]For more details on the ICC derivation in the linear programming framework, see Prescott and Townsend (1984). The key term is the 'likelihood ratio', $\frac{P(q|\hat{z},k)}{P(q|\bar{z},k)}$ which reflects the fact that by deviating the agent changes the probability distribution of output.

have,

$$\sum_{T\times Q\times W'\times K'} \pi(\tau, q, \bar{z}, k', w'|w, k)[U(\tau + (1-\delta)k - k', \bar{z}) + \beta w'] \geq$$

$$\geq \sum_{T\times Q\times W'\times K'} \pi(\tau, q, \bar{z}, k', w'|w, k)\frac{P(q|\hat{z}, k)}{P(q|\bar{z}, k)}[U(\tau + (1-\delta)k - k', \hat{z}) + \beta w'] \tag{10}$$

Apart from the additional ICC constraint (10), the MH regime differs from the FI regime in the set of feasible promised utilities, W. In particular, the lowest possible promise under moral hazard is no longer the value $w^{FI}_{\min}$ from (5). Indeed, if the agent is assigned minimum consumption forever, he would not supply effort above the minimum possible. Thus, the feasible range of promised utilities, W for the MH regime is bounded by:

$$w^{MH}_{\min} = \tfrac{U(c_{\min}, z_{\min})}{1-\beta} \text{ and } w^{MH}_{\max} = \tfrac{U(c_{\max}, z_{\min})}{1-\beta} \tag{11}$$

The principal cannot promise a slightly higher consumption in exchange for much higher effort such that agent's utility falls below $w^{MH}_{\min}$ since this is not incentive compatible. If the agent does not follow the principal's recommendations but deviates to $z_{\min}$ the worst punishment he can receive is $c_{\min}$ forever.

The constrained-optimal contract in the moral hazard regime, $\pi^{MH}(\tau, q, z, k', w'|w, k)$ solves the linear program defined by (6)–(10). The contract features partial insurance and intertemporal tie-ins, i.e., it is not a repetition of the optimal one-period contract (Townsend, 1982).

Limited commitment

The third setting with endogenously incomplete financial markets we study assumes away private information but focuses on another friction often discussed in the consumption smoothing and investment literatures (e.g., Thomas and Worrall, 1994; Ligon et al., 2005 among many others) – *limited commitment* (LC). As in those papers, by 'limited commitment' we mean that the agent could potentially renege on the contract after observing his output realization and realizing the transfer (τ) he is supposed to give to others through the intermediary. Another possible interpretation of this, particularly relevant for developing economies, is a contract enforcement problem. The maximum penalty for the agent reneging is for him to be excluded from future credit or risk-sharing – i.e., the assumption is the agent goes to autarky forever.

Using the same approach as with the other financial regimes, we write the optimal contracting problem under limited commitment as a recursive linear programming problem. The state variables are once again the capital stock, $k \in K$ and promised utility, w. The bounds of the set of promised utilities, W are set to $w^{LC}_{\min}$ equal to the autarky value at $k_{\min}$ (see Section 2.2.1) and $w^{LC}_{\max} = w^{FI}_{\max}$. Given the agent's current state (k, w) the problem of the intermediary is given by

$$V(k, w) = \max_{\pi(\tau,q,z,w',k'|k,w)} \sum_{T\times Q\times Z\times K\times W} \pi(\tau, q, z, w', k'|k, w)[q - \tau + (1/R)V(k', w')]$$

subject to the promise-keeping constraint

$$\sum_{T\times Q\times Z\times W\times K}\pi(\tau,q,z,w',k'|w,k)[U(\tau+(1-\delta)k-k',z)+\beta w']=w,$$

the limited-commitment constraints which ensure that reneging on the contract does not occur in equilibrium (respecting our timing that effort z is decided before output q is realized),

$$\sum_{T\times K\times W}\pi(\tau,\bar{q},\bar{z},w',k'|w,k)[U(\tau+(1-\delta)k-k',\bar{z})+\beta w']\geq\Omega(k,\bar{q},\bar{z})\text{ for all }(\bar{q},\bar{z})\in Q\times Z$$

and subject to non-negativity $\pi(\tau,q,z,w',k'|w,k)\geq 0$, technological consistency and adding-up,

$$\sum_{T\times K\times W}\pi(\tau,\bar{q},\bar{z},k',w'|w,k)=P(\bar{q}|\bar{z},k)\sum_{T\times Q\times K\times W}\pi(\tau,q,\bar{z},k',w'|w,k)\text{ for all }(\bar{q},\bar{z})\in Q\times Z,$$

$$\sum_{T\times Q\times Z\times K\times W}\pi(\tau,q,z,k',w'|w,k)=1,$$

Above, $\Omega(k,q,z)$ denotes the present value of the agent going to autarky forever with his current output at hand q and capital k, which is defined as:

$$\Omega(k,q,z)\equiv\max_{k'\in K}\ \{U(q+(1-\delta)k-k',z)+\beta v(k')\}$$

where $v(k)$ is the autarky regime value function defined in Section 2.2.1.

3 Computation

3.1 Solution Methods

We solve the dynamic programs for all financial regimes numerically[12]. Specifically, we use the linear programming (LP) methods developed by Prescott and Townsend (1984), Phelan and Townsend (1991) and Paulson et al. (2006). An alternative to the LP methodology in the literature is the 'first order approach' (Rogerson, 1985), used for instance by Abraham and Pavoni (2008), whereby the incentive constraints are replaced by their first order conditions[13]. A problem with that approach arises due to non-convexities introduced by the incentive and/or truth-telling constraints[14]. In contrast, the linear programming approach is extremely general and can be applied for *any* possible preference and technology specifications since, by construction, it convexifies the problem by allowing all possible lotteries over allocations. The only potential downside is that the LP method may suffer from the

[12] Given our primarily empirical objectives, we chose general functional forms that preclude analytical tractability. We verify robustness by using different parameterizations and model specifications.

[13] The first order approach requires imposing (strong) monotonicity and/or convexity assumptions on the technology (Rogerson 1985; Jewitt, 1988) or, alternatively, as in Abraham and Pavoni (2008), employing a numerical verification procedure to check its validity for the particular problem at hand.

[14] We do find such non-convexities in our solutions for the MH and UI regimes and hence we cannot use the first order approach as it is not always valid in our set up.

'curse of dimensionality'. However, as shown above, by judicious formulation of the linear programs, this deficiency is minimized. The main reason for using discrete grids for all variables is not the dynamic programming part, which can be also solved without discretization (e.g., using splines), but our linear programming approach to the MH and UI regimes (necessitated by non-convexities) and our empirical application using the likelihood of the discretized joint distribution of the data.

To speed-up computation, we solve the dynamic programs for each regime using a 'policy function iteration' algorithm (e.g., see Judd, 1998). That is, we start with an initial guess for the value function, obtain the optimal policy function and compute the new value function that would occur if the computed policy function were used forever. We iterate on this procedure until convergence. At each iteration step, that is for each interim value function iterate, we solve a linear program in the policy variables π for each possible value of the state variables.[15] In the unobserved investment (UI) regime the promised utilities set, $\mathbf{W}$ is endogenously determined and is solved for together with V. Using the incentive compatibility constraints, we restrict attention to non-decreasing promise vectors $\mathbf{w}(k)$. Specifically, we 'discretize' the set $\mathbf{W}$ by starting with a large set $\mathbf{W}_0$ consisting of linear functions $\mathbf{w}(k)$ with intercepts that take values from the grid $W = \{w_{\min}, w_2, ...w_{\max}\}$ defined in (11), and a discrete set of non-negative slopes. We initially iterate on the UI dynamic program using value function iteration, that is, we iterate over the promise set $\mathbf{W}$ together with the value function V, dropping all infeasible vectors $\mathbf{w}$ at each iteration and 'shrinking' $\mathbf{W}$ as a result (Abreu, Pierce and Stacchetti, 1990). Once we have successively eliminated all vectors in $\mathbf{W}$ for which the respective linear programs have no feasible solution, that is, once we have converged to the self-generating feasible promise set $\mathbf{W}^*$, we switch to (the much faster) policy function iteration and continue iterating on the Bellman equation until convergence[16]. The same approach is used for the set $\mathbf{W}_m$.

3.2 Functional Forms and Grids

Below are the functional forms we adopt for the empirical analysis. They are chosen and demonstrated below to be flexible enough to generate significant and statistically distinguishable differences across the financial regimes. Nevertheless, as argued earlier, our methods allow for any alternative or more general specifications of preferences and technology.

Agent preferences are of the CES form:[17]

$$U(c, z) = \frac{c^{1-\sigma}}{1-\sigma} - z^{\theta}$$

The production function, $P(q|z, k)$ represents the probability of obtaining output level, $q \in Q \equiv \{q_1, q_2, ..q_{\#Q}\}$, from effort $z \in Z$ and capital $k \in K$. In our baseline runs with Thai data we calibrate

[15] The coefficient matrices of the objective function and the constraints are created in Matlab while all linear programs are solved using the comercial LP solver CPLEX version 8.1. The computations were performed on a dual-core, 2.2 Ghz, 2GB RAM machine.

[16] We also verified the results against proceeding with value function iteration all the way.

[17] Our linear programming solution methodology does not require separable preferences. However, assuming separability is a common specification in the dynamic contracts literature so we adopt it for comparison purposes.

this function from a subset of households in our sample for which we have labor hours data.[18]

To get an idea of the computational complexity of the dynamic contracting problems we solve, Table 1 shows the number of variables, constraints, and linear programs that need to be solved at each iteration for each regime for the grids we use in the empirical implementation. The number of linear programs is closely related to the grid size of the state variables while the total number of variables and constraints depends on the product of all grid dimensions. The biggest computational difficulties arise from increasing $\#K$ or $\#Z$ as this causes an exponential increase in the number of variables and/or constraints. This is why we keep these dimensions relatively low, whereas using larger $\#T$ (or, equivalently $\#C$) is relatively 'cheap' computationally. In the unobserved capital regime (see Appendix A) the biggest computational difficulties arise from the huge number of linear programs to be computed over the two sub-stages.

The grids that we use in the estimation runs reflect the relative magnitudes and ranges of the variables in the Thai data or simulated data. In our baseline estimation runs we use a five-point capital grid K with grid points corresponding to the 10th, 30th, 50th, 70th and 90th percentiles in the data. The same applies for the output grid Q. We can use (and do robustness runs with) finer grids but the associated computational time cost is extremely high at the estimation stage because of the need to compute the linear programs and iterate at each parameter vector during the estimation. This is why we keep dimensions relatively low at present. Unfortunately, because of the extreme dimensionality and computational time requirements of the UI regime (see Table 1), we are currently unable to estimate it, with the exception of runs with coarse grids.[19] Table 2 displays the actual grids we use in the estimation runs with Thai rural data. Specific details on how all the grids are determined are given in Section 5.

4 Empirical Method

In this section we describe our estimation strategy. We estimate via simulated maximum likelihood each of the alternative dynamic models of financial constraints developed in Section 2. Our basic empirical method is as follows. We write down a likelihood function that measures the goodness-of-fit between the data and each of the alternative model regimes. We then use the maximized likelihood value for each model (at the MLE estimates for the parameters) and perform a formal statistical test (Vuong, 1989) about whether we can statistically distinguish between each pair of models relative to the data. We thus approach the data as if agnostic about which theoretical model fits them best and let the data themselves determine this. The results of the Vuong test, a sort of 'horse race' among competing models, inform us which theory(ies) fits the data best and also which theories can be rejected in view of the observed data.

[18] In robustness runs we also use a parametric form for $P(q|z,k)$ with parameters we estimate which encompasses production technologies ranging from perfect substitutes to Cobb-Douglas to Leontief forms.

[19] Currently, a single functional evaluation of the UI regime likelihood for our baseline grids takes about 45 minutes (as opposed to 7-9 sec in the MH regime) and over 1,500 such evaluations (47 days) are typically required to find the MLE parameters for a single estimation run. We are working on a parallel computing version of our estimation algorithm as well as optimizations based on the NPL approach (Aguirregabiria and Mira, 2002; Kasahara and Shimotsu, 2009).

The financial regimes we study have implications for both the transitional dynamics and long-run distributions of variables such as consumption, assets, investment, etc. Given our application to Thailand – an emerging, rapidly developing economy, we take the view that the actual data is more likely to correspond to a transition than to a steady state. Thus, estimating initial conditions (captured by the initial state variables distribution) is important for us, as well as fitting the subsequent trajectories using intertemporal data. In addition, our simulations show slow dynamics for the model variables in the mechanism design regimes and, theoretically, some regimes can have degenerate long-run distributions (e.g., immiserization in MH). These are further reasons to focus on transitions instead of steady states in the estimation. We explain the details below.

4.1 Structural maximum likelihood estimation

Suppose we have i.i.d. data $\{\hat{y}_j\}_{j=1}^n$ where $j = 1, ...n$ denotes sample units (in our application, households observed over seven years). For each j, the vector $\hat{y}_j$ can consist of different variables from a cross-section (e.g., consumption, income, capital, investment) or, if panel data is available as we use here, from different time periods (e.g., consumption at $t = 0$ and at $t = 1$). For example, in this paper we use (various subsets of) data from rural Thai households running small businesses on their productive assets, consumption and income, $\{\hat{k}_{jt}, \hat{c}_{jt}, \hat{q}_{jt}\}$ where $t = 0, ..., 6$ corresponds to years 1999-05. See Section 5.1 for more details.

We assume that all available data may contain measurement error. Assume the measurement error is additive and distributed $N(0, (\gamma_{me} r(x))^2)$ where $r(x)$ denotes the range of the grid X for variable x, i.e. $r(x) = x_{\max} - x_{\min}$ where x is any of the variables of interest used in the estimation. The reasoning behind this formulation is that for computational time reasons we want to be as parsimonious with parameters as possible in the simulated likelihood routine while still allowing the measurement error variance to be commensurate with the different variables' ranges. In principle, much more complex versions of measurement error can be introduced at the cost of computing time. The parameter γ_{me} is estimated in the MLE routine

The list of steps below describe the algorithm used to construct the simulated likelihood function.

1. Model solution to probability distribution

For any possible value of the state variables in a given regime (e.g., k, w – the capital stock, and the promised utility value for the MH, FI or LC models) and structural parameters ϕ^s (preferences, β, σ and technology, $P(q|z, k)$) the model solution obtained from the respective linear program (see Section 2) is a discrete joint probability distribution. For example, for the MH model the solution consists of the probabilities $\pi(\tau, q, z, k', w'|w, k)$ over the grids T, Q, Z, K, W where primes denotes future-period states. From this joint distribution we easily obtain (by manipulating the π's and summing over not needed variables, see Appendix C) the joint probability distribution over any desired set of variables to be used in the estimation – for instance the cross-section $\{c_{j0}, q_{j0}\}_{j=1}^n$.

In general, let us denote this distribution, for model m by $g_0^m(y^1|s^1, s^2, \phi^s)$ where y^1 is a vector of non-state variable data being fitted, s^1 is the vector of observable state variables and s^2 is the vector of unobservable state variables for that model. (We have $s^1 \equiv k$ for all models while $s^2 \equiv w$ or $s^2 \equiv b$

or s^2 absent, depending on the estimated model). The unobservable states, s^2 are treated as sources of unobserved heterogeneity endogenous to the models.

2. Initialization and unobservable state variables

To map the solution of each model to the data, we need to initialize the state variables s^1, s^2. For the unobservable state variables, b and w we assume that they initially come from a parametric initial distribution $\Omega(\phi^d)$ (e.g. normal, mixture of normals) the parameters of which, ϕ^d will be estimated in the SMLE routine.[20] We first integrate the joint probability distribution $g_0^m(y^1|s^1, s^2; \phi^s)$ over the unobserved state variable (e.g., in the MH model this is the variable w on the grid W). This is done using the assumed parametric distribution for this variable discretized via a standard histogram function (see Appendix for definition) applied on the grid W. The result is the joint distribution $g_1^m(y^1|s^1; \phi^s, \phi^d)$. For instance, this could be the joint distribution of c, q over $C \times Q$ for each $k \in K$ as given by the MH model solution integrated over the unobserved state distribution. Note that the joint distribution g_1^m depends on both the structural parameters ϕ^s as well as on the unobserved state distributional parameters ϕ^d. Naturally, since the only state in autarky is observable (k) this step is not performed when we estimate the autarky model.

For the observed state variables s^1 (here k) we take actual data $\hat{s}_j^1$ (i.e., $\hat{k}_j$) and discretize it over the model grid S^1 (i.e., K) via histogram function. We call the resulting distribution $H(\hat{s}^1)$. If dynamic data is used in the MLE, the $\hat{s}^1$ data come from the initial period of data used. Within this step, we allow for the possibility that the actual s^1 (in our application, k) data contains measurement error as explained above. In practice, this means manipulating the theoretical joint distribution $g_1^m(y^1|s^1; \phi^s, \phi^d)$ to transform it into the distribution $g_2^m(y^1|s^1; \phi^s, \phi^d, \gamma_{me})$ which is the model-m predicted joint distribution of y^1 over the compound grid Y^1 (e.g., $y^1 = (c, q)$ over $C \times Q$) for each value of the state variable s^1 at parameters ϕ^s, ϕ^d and including measurement error parametrized by γ_{me} in s^1. Note, it is computationally prohibitive to re-compute the model at non-grid points for s^1.

Next, given the theoretical joint distribution allowing for measurement error, $g_2^m(y^1|s^1; \phi^s, \phi^d, \gamma_{me})$, use the actual discretized distribution $H(\hat{s}^1)$ (which is inclusive of measurement error) to obtain the joint distribution over the estimated variables y (the model analogue of the data $\hat{y}$) implied by model-m when initialized at data $H(\hat{s}^1)$,

$$f^m(y|H(\hat{s}^1), \phi) \tag{12}$$

where $\phi \equiv (\phi^s, \phi^d, \gamma_{me})$. Here, we can have either $y = y^1$ (if the s^1 variables are not used in the estimation, in which case f^m is simply $g_2^m(.)$ integrated over s^1 with the probabilities from $H(\hat{s}^1)$ or we can have $y = (y^1, s^1)$ in which case f^m is the joint distribution of (y^1, s^1) over the compound grid $Y^1 \times S^1$. For instance, in the next section we have runs with $y = (c, q)$ (corresponding to the former case) and also runs with $y = (k, i, q)$ where $i \equiv k' - (1 - \delta)k$ is investment (corresponding to the latter case).

3. Measurement error and the likelihood function

[20] In the baseline runs we assume that the unobserved state w in the MH, FI, LC models is distributed $N(\mu_w, \gamma_w^2)$ while the unobserved state b in the B and S models is distributed $N(\mu_b, \gamma_b^2)$. This assumption is not essential for our methodology and more general distributional assumptions can be incorporated at the computational cost of additional estimated parameters.

Let $\Phi(.|\mu,\sigma)$ denote the pdf of $N\left(\mu,\sigma^2\right)$. We now allow for Normal measurement error in the estimated non-state variables y^1. Given the assumed measurement error distribution, the likelihood of observing data point $\hat{y}^1_j$ (e.g., $(\hat{c}_j,\hat{q}_j)$) relative to any model grid point $y^1_h \in Y^1$, for any $h = 1,...,\#Y^1$ is:

$$\prod_{l=1}^{L}\Phi\left(\hat{y}_j^{1,l}|y_h^{1,l},\sigma^l\right) \tag{13}$$

where $l = 1,...,L$ indexes the variables in $\hat{y}^1$ and where $\sigma^l = \gamma_{me} r(y^{1,l})$ is the measurement error standard deviation for each variable, as explained earlier.

Focus on the case $y = y^1$ (the case $y = (y^1, s^1)$ is handled analogously but the algebra is a bit more cumbersome since for each j we need to condition on its particular s^1 value). Expression (13) implies that the likelihood of observing data vector $\hat{y}_j$ (consisting of L components indexed by l) for model m, at parameters ϕ and initial conditions $H(\hat{s}^1)$ is

$$F^m\left(\hat{y}_j|H(\hat{s}^1),\phi\right) = \sum_h f^m(y_h|H(\hat{s}^1),\phi)\prod_{l=1}^{L}\Phi\left(\hat{y}_j^{1,l}|y_h^{1,l},\sigma^l\right) \tag{14}$$

where we assume that measurement errors in all variables are independent from each other. We basically sum over all grid points $h = 1,...\#Y^1$, appropriately weighted by f^m, the likelihoods in (14). For example, for consumption and income cross-sectional data $\hat{y}_j = (\hat{c}_j,\hat{q}_j)$ we have $F^m\left(\hat{c}_j,\hat{y}_j|H(\hat{s}^1),\phi\right) = \sum_h f^m((c,q)_h|H(\hat{s}^1),\phi)\Phi\left(\hat{c}_j|c_h,\sigma^c\right)\Phi\left(\hat{y}_j|y_h,\sigma^q\right)$ where $(c,q)_h$ go over all elements of $C \times Q$, $h = 1,...,\#C\#Q$.

Multiplying (14) over sample units (households) and taking logs, the simulated log-likelihood of the data $\{\hat{y}_j\}_{j=1}^n$, conditional on $H(\hat{s}^1)$ and given parameters ϕ and measurement error in s^1 and y^1 is (normalized by n)

$$\Lambda_n^m\left(\phi|H(\hat{s}^1)\right) \equiv \frac{1}{n}\sum_{j=1}^{n}\ln F^m\left(\hat{y}_j|\phi,H(\hat{s}^1)\right). \tag{15}$$

The maximization in (15) over ϕ is performed by an optimization algorithm robust to local maxima.[21] Standard errors are computed via bootstrapping, repeatedly drawing with replacement from the data.

4.2 Testing and model selection

We follow Vuong (1989) to construct and compute an asymptotic test statistic that we use to distinguish across the alternative models using simulated or actual data. The Vuong test does not require that either of the compared models be correctly specified. The null hypothesis of the Vuong test, is that the two models are asymptotically equivalent relative to the true data generating process – that is, cannot be statistically distinguished from each other based on their 'distance' from the data (in KLIC sense). The Vuong test-statistic is normally distributed under the null hypothesis. If the null is rejected (i.e.,

[21] We first perform an extensive grid search over the parameter space to rule out local extrema and then use the Matlab global optimization routines `patternsearch` and `fminsearch` (polytope algorithm) to maximize the likelihood.

the Vuong Z-statistic is large enough in absolute value), we say that the higher likelihood model is closer to the data (in KLIC sense) than the other. The pairwise nature of the test allows us to obtain a complete ranking by likelihood of all models we study.

More formally, suppose the values of the estimation criterion function being minimized (i.e., minus the log-likelihood) for two competing models are given by $L_n^1(\hat{\phi}^1)$ and $L_n^2(\hat{\phi}^2)$ where n is the common sample size and $\hat{\phi}^1$ and $\hat{\phi}^2$ are the parameter estimates for the two models.[22] The pairwise nature of the test conveniently allows us to obtain a complete ranking by likelihood of all models we study. Define the "difference in lack-of fit" statistic:

$$T_n = n^{-1/2} \frac{\Lambda_n^1(\hat{\phi}^1) - \Lambda_n^2(\hat{\phi}^2)}{\hat{\sigma}_n}$$

where $\hat{\sigma}_n$ is a consistent estimate of the asymptotic variance, σ_n of the log of the likelihood ratio, $\Lambda_n^1(\hat{\phi}^1) - \Lambda_n^2(\hat{\phi}^2)$.[23] Under regularity conditions (see Vuong, 1989, pp. 309-13 for details), if the compared models are strictly non-nested, the test-statistic T_n is distributed $N(0,1)$ under the null hypothesis.

5 Application to Thai Data

5.1 Data

In this section we apply our estimation method to actual household-level data from a developing country. We use the Townsend Thai Data, both the rural Monthly Survey and the annual Urban Survey (Townsend, Paulson and Lee, 1997). Constrained by space and computation time, we report primarily on the rural data, but where possible compare and contrast results with the urban data.

The Monthly Survey data were gathered from 16 villages in four provinces, two in the relatively wealthy and industrializing Central region near Bangkok, Chacheongsao and Lopburi and two in the relatively poor, semi-arid Northeast, Buriram and Srisaket. That survey began in August 1998 with a comprehensive baseline questionnaire on an extensive set of topics, followed by interviews every month. Initially consumption data were gathered weekly, then bi-weekly. All variables were added up to produce annual numbers. The data we use here begins in January 1999 so that technique and questionnaire adjustments were essentially done. We use a balanced panel of 531 rural households who run small businesses observed for seven consecutive years, 1999 to 2005.

Consumption expenditures, c include owner-produced consumption (rice, fish, etc.). Income, q is measured on accrual basis (see Samphantharak and Townsend, 2009) though at an annual frequency this is close to cash flow from operations. Business assets, k include business and farm equipment, but exclude livestock and household assets such as durable goods. We do perform a robustness check with

[22] For the functional forms and parameter space we use in the estimation runs, the models we study are statistically non-nested. Formally, following Vuong (1989), if two models are represented by the parametric families of conditional distributions $\mathcal{F} = \{F_{Y|Z}(.,.|\phi^1) : \phi^1 \in \mathbb{R}^{d_{\phi^1}}\}$ and $\mathcal{G} = \{G_{Y|Z}(.,.|\phi^2) : \phi^2 \in \mathbb{R}^{d_{\phi^2}}\}$ where $\{Y_i, Z_i\}_{i=1}^n$ is i.i.d. data, they are non-nested if $\mathcal{F} \cap \mathcal{G} = \varnothing$. The Vuong test can be also used for overlapping models, i.e. neither strictly nested nor non-nested, in which case a two-step procedure is used (see Vuong, 1989, p. 321).

[23] In practice one uses the sample analogue of the variance of the LR statistic (see Vuong, 1989, p. 314).

respect to the asset definition — see Section 6.1. Assets other than land are depreciated. All data are in nominal terms but inflation was low over this period. The variables are not converted to per-capita terms, i.e., household size is not brought into consideration (though we do a robustness check below). We construct a measure of investment using the assets in two consecutive years as: $i_t \equiv k_{t+1} - (1-\delta)k_t$ for each household.

From the Urban Survey which began in November 2005, we use a balanced panel of 475 households observed each year in the period 2005 to 2009 from the same four provinces as in the rural data plus two more, Phrae province in North Thailand and Satun province in the South. We use the same variables — consumption expenditure, business assets and income as in the rural data

Table 3 displays summary statistics of these data in thousands Thai baht for both the rural and urban samples (the average exchange rate in the 1999-05 period was 1 USD = 41 Baht). All magnitudes are higher in the urban sample because those households are richer on average. Productive assets, k and investment, i are very unequally distributed as reflected in the high standard deviations and ratio between mean and median. There are many observations with zero of close to zero assets and few with quite large assets. More detail is reported in Samphantharak and Townsend (2011, ch. 7) for the monthly data.

Figure 1 plots, for both the Urban and Monthly data, deviations from the sample year averages for income, consumption, and investment and illustrates visually the degree of smoothing relative to income fluctuations. We see that there is significant degree of consumption smoothing in the data but smoothing is not perfect as the full insurance hypothesis would imply (if all households were identical, the right hand panel should be flat at zero). Investment should not move with cash flow fluctuations either (controlling for productivity), but should move with investment opportunities. This is more so in the urban data.

Figure 2 plots the relationships between consumption level changes and assets level changes each relative to income level changes, following Krueger and Perri (2010). We sort the data into twenty bins by average income change over the seven years and report the average consumption and assets change corresponding to each bin (each marker corresponds to a bin). The comovement between consumption changes and income changes Δc and Δq in the Figure can be viewed as a measure of the degree of consumption smoothing in the data. For the rural data, the top left panel shows a positive relationship between consumption and income changes (correlation of 0.11) but the slope of the consumption change line is smaller than that found in Krueger and Perri for Italian non-durable consumption data indicating more consumption smoothing. This pattern is a bit stronger, i.e., the line is flatter in the urban data (bottom left panel). The top right panel shows that changes in business assets in the rural data are positively correlated with changes in contemporaneous income although this general pattern does not hold at the two extremes of very high or very low income level changes (possibly due to measurement error or large outliers). The comovement between assets and income changes is stronger in the urban data (bottom right panel).

5.2 Structural estimation results

We convert all data from Thai currency into 'model units' by dividing all currency values by the 90-th percentile of the assets distribution in the sample (this is 179,172 Thai baht for the rural sample). The normalized asset values are used to define a five-point assets grid[24], K corresponding to the 10th, 30th, 50th, 70th and 90th percentiles in the data (e.g., for the rural sample we have $K = \{0, .02, .08, .33, 1\}$). The unequal spacing reflects the skewness of the asset distribution in the data. Similarly, from normalized income q we define a five-point grid corresponding to the 10th, 30th, 50th, 70th and 90th percentiles of q (for the rural sample, $Q = \{.04, .17, .36, .75, 1.75\}$). These grids imply an upper bound of .82 model units for the borrowing/savings B grid in the B regime to ensure no default. The consumption grid used to compute the moral hazard, full information and limited commitment models consists of thirty-one equally spaced points on the interval $[.001, .9]$ model units which covers the range of consumption expenditure in the data. In the other regimes consumption is a residual variable, not independently chosen but obtained from the k and b grids. The W grid is endogenously determined as described in Section 2.2. Table 2 contains the grids used in the baseline estimation runs.

We use the algorithm described in Section 3 to estimate each model. For each regime, we estimate the structural parameters σ (risk aversion) and θ (effort curvature), together with the distributional parameters, μ_w, γ_w (respectively, μ_b, γ_b for B or S), and the standard error parameter, γ_{me}. The discount factor β, the risk-free rate, R and the depreciation rate, δ are fixed at $\beta = 0.95$; $R = 1.05$ and $\delta = 0.05$. We also do robustness runs with alternative values for R and δ. For robustness purposes when using cross-sectional data we estimate using the first two or the final two years of the Thai data, for example the monthly 1999-00 data or the 2004-05 data. In the latter case, we act as if the data starts in 2004, i.e., the initial k distribution, $H(\hat{s}^1)$, see Section 4.1, is formed from 2004 data. We also perform runs in which we initialize the models with 1999 k data, run them for a number of periods and estimate with later periods' cross-sectional data.

Parameter estimates and bootstrap standard errors are displayed in section 5.2.1 and Table 4 (reported for the monthly 1999-00 data runs only, to save space). Then we turn to the model comparisons in section 5.2.2 and Table 5.

5.2.1 Parameter estimates

We report first the parameter estimates for the best fitting regimes. Table 4 reports the maximized likelihood of each model in the last column. The best-fitting model is denoted with an asterisk in the first column. The saving only (S) regime is the best-fitting (achieves the highest likelihood) with *kiq* or *cqik* data or tied for best-fit with the moral hazard (MH) model (second in likelihood) with *cq* data. The parameter estimates for the S regime show that the estimate for γ_{me} – the relative measurement error size parameter is relatively low (in the range 0.09-0.13 depending on the data used). This corresponds to measurement error with standard deviation of 9 to 13% of the variables range. The estimates for the risk aversion parameter σ indicate relatively high degree of curvature in consumption (2.96 when using the *cq* data and 5.7 with *kiq* and *cqik* data with the exception of .5 for MH with *cq* data). The

[24] We use a standard histogram function based on distance to the closest grid point (Matlab's command `hist`).

parameter θ can be interpreted as capturing the extent to which households dislike variability in effort is estimated to be relatively high (9.2) when using *kiq* data alone, but low (0.10-0.11) when including consumption data, in the *cqik* and *cq* data runs. Since effort z takes values on $(0, 1)$, values for θ close to zero imply relatively high effort disutility, z^θ over the whole effort range, while high values of θ such as 9.2 imply low effort disutility at low z levels increasing sharply at high effort levels. The final two parameters reported in Table 4, $\mu_{w/b}$ and $\gamma_{w/b}$ determine the mean and standard deviation of the initial distribution of unobserved heterogeneity in the models captured by the state variables 'promised utility', w (for the MH, FI and LC models) and debt/savings, b (for the B and S models). Their estimates are reported relative to these variables' grid ranges, e.g. $\mu_w = 0.55$ estimated for the best-fitting MH regime with *cq* data means that the initial promised utility distribution is estimated to have a mean of $w_{\min} + 0.55(w_{\max} - w_{\min})$. Similarly, γ_w refers to the estimate of the initial distribution standard deviation, again relative to the grid range (e.g., $\gamma_w = .05$ means a standard deviation of $.05(w_{\max} - w_{\min})$ around the mean μ_w). The estimates of μ_b for the S regime are high and in a very narrow range (0.96-0.99) independently of the data used which means putting the mean of the unobserved savings distribution close to zero savings (the upper limit of the B grid). The standard deviation parameter γ_b is also tightly estimated and relatively small, in the range 0.01-0.08 depending on the data used.

More generally, Table 4 shows the estimates differ across the regimes as the MLE optimization adjusts the parameters for each model to attain best fit with the Thai data. The estimates for the measurement error size, γ_{me} across all estimated regimes and data types are in a relatively narrow range 0.09-0.18 (with one exception, LC with *cqik* data). This corresponds to measurement error with standard deviation of 9 to 18% of the variables grid span. The best-fitting model in each sub-section of Table 4 accommodates the smallest (with *kiq* and *cqik* data) or close to the smallest (with *cq* data) measurement error size. In contrast, the regimes which obtain the lowest likelihoods generally feature higher estimated level of measurement error (e.g., the MH, FI, LC regimes with *kiq* data have estimates for γ_{me} larger than 0.148 vs. 0.089 for S). The likely explanation is that, to compensate for the bad fit, the MLE procedure is raising the level of measurement error. The estimates for the risk-aversion parameter σ vary in the range 0 to 5.7 depending on the model and data used. The MH and FI regimes seem to require less risk-aversion to fit the data (σ below 0.65) while the B, S and LC models produce estimates above 2.5 (with one exception, B with *cqik* data). One possible intuition is that the MH and FI models need to impose less curvature in consumption to explain the relative lack of smoothing in the data – higher curvature will lead to 'excessive' smoothing in the model relative to the data. There is substantial variation in the estimates for the effort-curvature parameter θ across the financial regimes. For example, with the joint *cqik* data, it ranges from 0.11 for the S regime to 9.2 for the autarky model. Table 4 shows that, in order to fit the data, the B and S regimes put the mean of the unobserved debt/savings distribution in the range 0.84 to 0.98 which is close to the borrowing limit $b_{\max} > 0$ in the B model (or to zero savings for the S model). The variance parameter ($\gamma_{b/w}$) is estimated to be relatively low (below 0.1 of the w or b grid ranges) with the LC, *cq* data case the sole exception.

5.2.2 Model comparisons – Vuong test results

Business assets, investment and income data

We first explore the implications of the different regimes on the production side alone by using the joint distribution of assets, investment and income, *kiq* in the data. When estimated from *kiq* 1999/00 cross-sectional data, the financial regimes rank in decreasing order of likelihood as: S, B, A, MH, FI, LC (Table 4, last column). The saving only (S) model wins all its bilateral model comparisons at the 1% significance level in the 1999-00 data (Table 5, row 1.1) while the B and S regimes are tied for best fit in the 2004/05 data (row 1.2). The MH, LC, FI and A regimes obtain the lowest likelihoods and are rejected at the 1% significance level in all comparisons with the best fitting regimes.

Consumption and income data

We next test whether we can distinguish between the regimes based solely on the degree of consumption smoothing they imply relative to the data using the *cq* joint cross-sectional distribution (Table 5, section 2). The Thai consumption and income data alone seem to be unable to pin down precisely the best fitting regime — we obtain a tie between moral hazard and savings only.

Business assets, investment, consumption and income data

We also evaluate the gains from using combined data on assets, income, and consumption as opposed to using *kiq* or *cq* data on their own. This captures both the consumption and production side of the household/enterprise problem as shaped by the financial constraints. Theoretically, with incomplete markets the classical separation between consumption and production/investment decisions fails, so empirically, joint data on the consumption and investment side of the model should be helpful to distinguish the alternative regimes better (section 3 of Table 5). Adding the consumption data to the set of variables, the joint distribution of which we estimate, we observe an improvement in our ability to distinguish the regimes relative to using *kiq* or (especially, see the decrease in the number of ties) *cq* data alone and we are able to identify S as the single winner. The limited commitment (LC) regime does not fit the data well.

Dynamics and panel data

We also estimate and test across the alternative regimes employing the panel data, targeting differences in the variables' dynamics across the models. Our empirical method based as it is on computing the joint distribution of variables of interest does not allow us to use more than two years of the data at a time; computing joint probabilities of the model variables over time becomes computationally intensive and the relatively small sample size we have in this application does not allow to precisely map the joint distribution of *cq*, *kiq* or *cqik* over many periods. Since we require i.i.d. data to form the likelihood we also cannot simply pool observations from different time periods. For example, if we wanted to use three periods of *kiq* data in the MLE, we would have a discretized joint distribution f^m in (12) of dimension (total number of mutually exclusive probability cells) equal to $(\#K)^3(\#I)^3(\#Q)^3$ which (see Table 2) greatly exceeds the number of observations we have. For those reasons we only use c, q data for two periods below.

Restricting ourselves to two years (section 4 in Table 5), the regimes' likelihood ordering is consistent with results discussed above. The ability to distinguish across the regimes with the 1999/00 *cq* panel data is no worse than in the single 1999 cq cross-section but the ability to distinguish the regimes worsens with the gap between years included in the panel, with three regimes, now including the borrowing regime, statistically tied with MH and S when the 1999/05 cq panel is used.

Second, we did runs (Table 5, section 5) where we used the 1999 distribution of k to initialize the models but 2004-05 data in the MLE. That is, each model was run for six (seven in the case of cq data) periods and its predictions for the joint distribution of k, i, q, c, q or c, q, i, k at that period were matched with actual data from the corresponding period of the panel. Once again, the evidence points to the exogenously incomplete market regimes fitting best (B, followed by S) when using the production-side (kiq) or the joint ($cqik$) data, while we are unable to distinguish those two regimes from FI and MH with the cq data alone. We do discuss further the models' ability to match the dynamics of the data in the robustness Section 6.

Rural networks

In Table 6 we also perform several runs with sub-samples of the Thai rural data with various definitions of networks. We first focus on a set of households (n = 391) who are related by blood or marriage (section 1 of Table 6). Compared to the whole sample results and likelihoods (Table 5), with consumption and income cross-sectional data alone, this networked sub-sample of households allows us to pin down unequivocally the best fitting regime as moral hazard. Evidently, family networks help in consumption smoothing as in Chiappori et al. (2010). We are unable to pin down the regime in the cq sub-sample of unrelated households (all regimes are tied) due to small number of observations. In contrast, including the production side, the 1999-00 kiq and $cqik$ data still indicate that the S regime remains best-fitting with all stratifications, as in the whole sample. This is also true in the sub-sample of households not related by blood or marriage (line 1.5).

A re-estimation with another data sub-sample, of 357 households related via observed personal loans or gifts (Kinnan and Townsend, 2009) and 1999 cq data puts the FI regime on top in terms of likelihood (row 2.1 in Table 6), compared to less information for those not in networks, but again this effect is not present when using the investment and capital data in the same sub-sample — S wins again.

Urban data

We also ran our estimation routines on the more recent panel data from the Thai Urban Surveys on households in urban areas and displayed in Table 7. Now the moral hazard (MH) regime is identified as the best-fitting using the complete data on consumption, income, investment and assets (c,q,i,k) from both 2005/06 and 2008/09 (see Table 7, section 1).

In fact, the moral hazard regime appears as winning in every row of the table with one exception. The exception is when using data on the production side alone, as in the (k,i,q) specification (Table 7, section 3), where we recover the saving only (S) financial regime, the same as in the rural data. But again with joint production and consumption data (c,q,i,k), the results are more decisive and an endogenously incomplete (MH) rather than exogenously incomplete regime is chosen. Urban households seem to be doing better in terms of smoothing consumption and investment. We come back to this in Section 7.

6 Robustness

6.1 Additional estimation runs with Thai data

We perform a large number of additional estimation runs to check the robustness of our baseline results and shed more light onto the regime likelihood patterns with consumption vs. investment vs. joint data. Here we restrict our robustness checks to the rural data.

Using estimated production function

We re-do the baseline estimation runs from Table 5 using an alternative, estimated production function the parameters of which are estimated jointly with the other model parameters. The advantage to this is that we avoid endogeneity in production decisions, determined jointly with the financial regime. The disadvantage is that we are then jointly testing the regime with the assumed parametric form. We assume the following parametric form for $P(q|z,k)$,

$$P(q = q_1|z,k) = 1 - (\eta k^\rho + (1-\eta)z^\rho)^{1/\rho} \tag{16}$$
$$P(q = q_i|z,k) = \frac{1}{\#Q - 1}(\eta k^\rho + (1-\eta)z^\rho)^{1/\rho} \text{ for } i = 2,..,\#Q$$

where q_1 is the lowest output level. The probability of obtaining each output level is bounded away from zero. We estimate ρ but fix $\eta = 1/2$ for computational time reasons. This functional form encompasses a range of production technologies. ($\rho = 1$ – perfect substitutes technology; $\rho \to 0$ – Cobb-Douglas; and $\rho \to -\infty$ – Leontief). Note that P determines expected and not actual output and so the CES parameter ρ here is not comparable to values from the macro literature.

Table 8 reports the results from the Vuong model comparison tests. Using the 'production side' of the household operation, (k,i,q) from either 1999/00 or 2004/05 (Table 8, section 1) reveals the saving only (S) regime as best fitting, consistent with the baseline runs (Table 5, section 1) though one tie is now eliminated. Using the 'consumption side' the B regime comes out on top tied with S and MH, the best fitting regimes in Table 5. The combined consumption, income, investment and assets data runs (Table 8, section 3) likewise elevate B to a tie with S, the dominant regime in Table 5. The results with two-year panels of (c,q) data (Table 8, section 4) are quite similar to those in Table 5, with many ties. The bottom line is that most results are similar, and so we conclude that assumptions about the production function are not driving the regime comparisons.

Robustness runs – risk neutrality, fixed measurement error, adjustment costs, removed fixed effects, etc.

We perform a large number of robustness check with various versions of our baseline specification. The results are displayed in Table 9. Unless stated otherwise Table 9 uses 1999-00 data. First, a re-estimation imposing risk neutrality, i.e., fixing $\sigma = 0$ instead of estimating σ, (Table 9, section 1) produces similar results to the baseline runs allowing for risk aversion. One minor difference is that the MH regime shows as the sole winner using the 1999 Thai (c,q) data (row 1.1 in Table 9). Otherwise, imposing risk neutrality pins down the borrowing regime, alone or tied with saving only.

Re-estimating with fixed magnitude of measurement error (in these runs we set the standard deviation parameter γ_{me} to 0.1; see Table 9, section 2) naturally reduces the regimes' likelihood values (especially for autarky) but preserves the MH and S models' best fit with the Thai data on consumption and income (although now tied with FI too). The saving only (S) regime emerges as the single best-fitting regime with (k,i,q) and (c,q,i,k) data, as in the baseline runs.

Section 3 of Table 9 allows quadratic adjustment costs in investment. The bottom line here is that the introduction of adjustment costs does blur the distinction across the financial regimes, especially, endogenous versus exogenously incomplete regimes, or can pick out a different regime entirely. The full information regime with adjustment costs corresponds to the standard adjustment costs model in the literature (Bond and Meghir, 1994 among many others) though we are allowing risk aversion. It appears as best-fitting in a tie with S using joint production and consumption data (line 3.3). Production data alone (k,i,q) actually has S tied with autarky, A (line 3.2). The borrowing regime with adjustment costs fits the consumption-income data best (line 3.1), though it was dominated when adjustment costs were not included, as in Table 5. It is true that the likelihoods of some the regimes improves, but then again, there are more parameters. We do not come away convinced that adjustment costs offer a better underlying specification. An exception, dealing with the persistence of capital, is discussed below.

In section 4 of table 9 we perform several estimation runs with the 1999/00 data from which fixed effects have been removed and replaced with average values. The worry here is that there may be more heterogeneity in the raw data than what the model is designed to accommodate. Removing only year fixed effects from the (c,q,i,k) data (row 4.1) reassuringly produces identical result as the corresponding baseline run (row 3.1 in Table 5). Removing both year and household fixed effects (row 4.2) reveals S as the best-fitting regime with the (k,i,q) data, also as in the baseline. The MH regime has the highest likelihood with 1999 consumption and income data from which year and household fixed effects have been cleaned (row 4.3), which is again consistent with the baseline where MH is tied with S. The run with (c,q,i,k) data where MH is revealed as best-fitting (row 4.4 in Table 9) is the only one that does not match the baseline (where S is found to be best-fitting, see row 3.1 in Table 5). Naturally, removing time and household fixed effects from the Thai data and replacing them with averages produces a 'smoother' dataset so perhaps the better fit of the MH model is not surprising. We come back to this issue in the estimation runs we do with simulated data from one of the models (see Section 6.2 below). In one of these runs we deliberately generate simulated data with heterogeneity in either risk aversion or productivity and estimate the models as if this heterogeneity did not exist. In another run we took out the heterogeneity the model is generating endogenously and the results echo what we find above.

In section 5 of Table 9 we perform several additional robustness runs using 1999-00 (c,q,i,k) data. We check robustness with respect to our definitions of assets and income in the Thai data (Table 9, row 5.1). We re-estimate using the (c,q,i,k) data including all household assets and livestock in the definition of k and exclude households who have only labor income. The sample size drops to 297 but our main findings about the best fitting regime from Table 5, row 3.1 do not change. Re-estimating fixing the lender risk-free rate to $R = 1.1$ instead of 1.05 in the baseline or the depreciation rate $\delta = 0.1$ instead of 0.05 in the baseline (Table 9, rows 5.2 and 5.3) does not affect our findings either

– the S regime is still best-fitting although in these runs we cannot discern it from the B regime. A re-estimation using either coarser grids than the baseline (3 points) or denser grids (10 points) produces similar results.

Finally, section 6 of Table 9 reports results from computing two additional financial regimes – a moral hazard regime with unobserved investment, UI (see Appendix A for detailed description) and a regime with hidden output, HO (see Appendix B for detailed description). We do this to show the generality of our approach and its ability to accommodate various models of information constraints, including those in the literature reviewed in the introduction, outside our six baseline regimes. Due to extreme computational time requirements we are unable to estimate the UI regime for all runs in Tables 5-8 and we had to use the coarse, three-point grid specification (read together with line 5.4 in Table 9). The UI regime outperforms the other mechanism design regimes (MH, FI and LC) in likelihood. The UI and HO regimes achieve worse fit with the data compared to the S and B regimes so our overall conclusions from the baseline runs stand.

We also tried runs with data stratified by region and by 'net wealth' constructed from the household accounts as in Samphantharak and Townsend (2011). We did not find systematic patterns or differences with the baseline.

6.2 MLE with simulated data

Because of the analytical complexity of the dynamic incomplete markets models we study, it is not possible to provide theoretical identification proofs while keeping the setting sufficiently general. We are aware (e.g., Honore and Tamer, 2006) that point identification may sometimes fail in complex structural models like ours. To try address this possibility we use a form of 'numerical identification' algorithm consisting of the following steps: Step 1 – take a 'baseline' model parametrized by a vector of parameters, ϕ^{base}; Step 2 – generate simulated data from the baseline model; Step 3 – estimate the baseline model using the data from Step 2 using our method described in Section 4 (with variable grids determined by the simulated data percentiles and using the same parameter grid search and global maximization routines as we use in the Thai data) and obtain estimates, $\hat{\phi}^{base}$; Step 4 – if the estimates from Step 3 are numerically close to the baseline ϕ^{base} within the standard error bands and the Vuong tests recover the actual data-generating regime, report success, otherwise report failure. In other words, we use data simulated from the model itself to verify that our estimation methodology performs as it should.

Specifically, we adopt as a baseline the moral hazard regime (MH) and simulate an artificial panel dataset from it of size $n = 1000$, which we then use to estimate and test across all regimes, including MH. Details on how the data were simulated are in Appendix D. All runs in this section use the specification with parametric production function – see (16) in Section 6.1.

The results reported in this section are representative of many more runs that we did, including several alternative parametrizations. We discuss some of these in Section 6.2.3 though obviously we are limited by space in what we can report. While we have done our best to demonstrate the robustness of the methods, this section is primarily intended as a 'validation' exercise of our empirical method to support the results with the actual Thai data.

6.2.1 Baseline results

For each model we follow the procedures described in Section 3 and 4 and Appendix D— we compute the model, generate an initial state distribution, then generate simulated data adding on measurement error (two specifications, with high and low variance), and finally compute the joint probability distribution of the variables of interest that enter the likelihood function. In particular, we select new variable grids based on the percentiles of the simulated data as in Section 5.2 — those grids do not coincide with the ones used to generate the data as if we are applying our method agnostically to a dataset we know nothing about. One of the estimated models is the data-generating regime (MH) itself in order to verify whether we can recover the data-generating parameters (see the discussion in section 4.3). Finally, we perform Vuong tests to establish whether we can distinguish statistically between the data-generating (here, MH) and the rest of the regimes, as well as between any counterfactual regime pairs (e.g., B and S). The description of results below follows the layout of Section 5.2.

Business assets, investment and income

We first estimate and test the financial regimes based on their implications about assets, investment and income. To that purpose we simulate a data sample of size $n = 1,000$ from the joint distribution of (k,i,q) in the baseline (MH) regime. Tables 10 (for the low-measurement error case only) and 11 (section 1) display respectively the parameter estimates and Vuong test results with these data. Table 10 shows that, when estimating the data-generating MH regime with (k,i,q) data the baseline parameter values (last row in each section, in italics) used to generate the data for γ_{me}, σ and ρ are recovered relatively well given the ranges on which they take values but θ, μ_w and γ_w are off. In terms of likelihoods, the MH and FI models achieve the highest likelihood, followed by the S and B regimes, LC and finally autarky. The parameter estimates in Table 10 differ across the estimated regimes, as the MLE procedure is trying to fit the common data, but the estimates are quite similar between the FI and MH regimes estimated with (k,i,q) data. This is not the case for the other models where to fit the data some of the parameters (e.g. ρ) take values far from the data-generating ones. The autarky regime requires very high measurement error to fit the data compared to the baseline value for γ_{me} (0.1).

Turning to the Vuong tests (Table 11, section 1), we find that in the low measurement error specification ($\gamma_{me} = 0.1$) we are able to distinguish between the data-generating regime (MH) and the LC, B, S and A regimes at the 1% significance level except with (k,i,q) data for which we cannot discern the MH and FI regimes statistically. We also distinguish across the regimes in all pairwise comparisons between counterfactual (non-MH) regimes. That is, even if the researcher (incorrectly) believes that the data were, for example, generated from the complete markets FI regime, he/she can still distinguish it from the exogenously incomplete B, S and A regimes. In contrast, with high measurement error ($\gamma_{me} = 0.2$) in the baseline simulated data, the distinction between the regimes is blurred and, based on the Vuong test, we cannot discern statistically between moral hazard and all regimes but autarky. In all cases, including with high measurement error, all non-autarky regimes are statistically distinguishable at the 1% level from the autarky regime.

Consumption and income

We next estimate with data only on the consumption side using the (c, q) joint distribution. The likelihood values (see Table 10) are ordered MH, FI, S, LC, B and A from highest to lowest. Table 10 shows that using (c, q) data we recover the data-generating parameters better than with (k, i, q) data with the exception of ρ (compare the row for the MH regime estimates with the baseline parameters row for in Table 10). The parameter estimates for the exogenously incomplete regimes in many instances differ significantly from the baseline parameters used to generate the data.

With low measurement error, the baseline (MH) regime is distinguished at the 1% significance level from all alternatives (Table 11, section 2). However, ties appear between some counterfactual pairs (FI and S, LC and B) or lower significance values. With high measurement error we are, however, unable to recover the MH regime as best-fitting. The autarky regime is once again statistically distinguished from all others, including in the high measurement error specification.

Business assets, consumption, investment, and income

The parameter estimates with (c,q,i,k) data are reported in the third section of Table 10. The data-generating regime is recovered as best-fitting and the parameter estimates for the MH regime are very close to the baseline parameters (compare with the (k,i,q) case in particular, especially θ, $\mu_{w/b}, \gamma_{w/b}$). The regime order in likelihood is MH, FI, LC, B, S, A; and the 'incorrect' regimes require higher measurement error to fit the data compared to the .1 baseline.

Section 3 of Table 11 reports the estimated likelihoods and Vuong test results using (c,q,i,k) data from the MH regime. The ability to distinguish the data generating regime from all alternatives is perfect (at the 1% level) with low measurement error (and only tied with FI with high measurement error). These results (compare with the corresponding results for cq or kiq data, especially for the high measurement error specification) clearly demonstrate that using the joint data on consumption and production yields a significant improvement. The ability to distinguish between counterfactual regimes (those different from MH) also improves significantly relative to when using (c,q) data or (k,i,q) data alone – the number of ties falls from seven or ten to two overall.

Dynamics and panel data

As in Section 5 we also estimate and test the models using their implications about the dynamics of consumption and income. Specifically, we use simulated data on the joint distribution of consumption and income, (c,q) in two different periods: $t = 0$ and 1 (see rows 4.1 and 4.2 in Table 11) or $t = 0$ and 50 (see row 5.1) as in a panel dataset. Compared to Table 11, section 2 which uses a single (c,q) data cross-section, section 4 of Table 11 demonstrates that using intertemporal data significantly improves our ability to distinguish the regimes (both data-generating and counterfactual) – the number of ties diminishes from two or seven – see rows 2.1 and 2.2 – to zero or three ties, depending on the panel time span and measurement error. The improvement in ability to discern the regimes using intertemporal (c,q) data is comparable to when the joint data on (c,q,i,k) was used, especially with low measurement error (compare rows 3.1-5.1 with 1.1-2.1).

6.2.2 Additional results with simulated data

Measurement error size, sample size, grid size

Table 11, section 5 contains results from several additional robustness runs using (c, q, i, k) data generated from the MH regime in the low measurement error specification unless stated otherwise. Row 5.2 analyzes the effect of generating the data without adding measurement error (we do allow measurement error in the MLE, when estimating afterwards). The Vuong test statistics show that, as in line 3.1, the data-generating MH model is distinguished at the 1% confidence level from the rest of the regimes but some counterfactual regime pairs are indistinguishable. The autarky regime remains statistically distinguishable from all others. Rows 5.3 and 5.4 of Table 11 study the effect of varying the simulated data sample size. We find that reducing the sample size, n from 1,000 to 200 produces more ties between counterfactual regimes (compare with line 3.1) but the MH regime is still distinguished at 1% confidence level. Increasing the sample size of simulated data to 5,000 achieves virtually the same results to the $n = 1,000$ baseline run with (c, q, i, k) data (compare lines 5.4 and 3.1). Row 5.5 of Table 11 checks the sensitivity of our results to grid dimensionality. Reducing the size of all grids to three points does not affect the Vuong statistics and baseline results in any significant way relative to row 3.1 which is reassuring. In row 5.6 we used denser grids for K and W (10 points) to generate the simulated data but then estimated with the baseline five-point grids obtained from the data percentiles as explained above. We see that the results are once again not sensitive to the grid sizes.

Allowing for heterogeneity in productivity

We next test the robustness of the baseline results by allowing for 'productivity' differences across households. Remember that in contrast, in our baseline model any initial heterogeneity across agents can exist only in their assets (k), debt/savings (b) or 'promised utility' (w) endowments. Specifically, we draw ten productivity values from a uniform distribution on [0.75,1.25] and compute the MH regime multiplying the grid Q by each productivity factor, to capture 'skill' heterogeneity. We draw the simulated data from these heterogeneous joint distributions, ending up with (c, q, i, k) data that corresponds to that of a mixture of households with different productivities. We then estimate all six regimes as if those differences in productivity did not exist (that is, as if we mistakenly treat the data as generated without such differences, as we do in the runs with actual data in Section 5). Line 5.7 in Table 11 shows that allowing for this additional source of unobserved heterogeneity (and thus, mis-specification) in our model does not affect the robustness of the baseline results. Significantly, we still recover MH as the best-fitting regime, distinguished at the 1% level from all others. Relative to the homogeneous skill baseline (line 3.1) some ties between counterfactual regimes do appear. The parameter estimates are also quite close to the baseline (see Table 10), for the MH regime we obtain γ_{me}=.103, σ=.5, θ =2.36, ρ=.26, μ_w=.51 and γ_w=.34.

Allowing for heterogeneity in risk aversion

We also test the robustness of our method to heterogeneity in preferences in the data with which we estimate. Namely, we generate data from the MH regime at three different values for the risk aversion parameter σ (holding all other parameters at their baseline values), $\sigma = 0.62$, 0.78 and 1.4 taken from the range of risk aversion values estimated in Schulhofer-Wohl and Townsend (2011). Again, a 'mixed' data sample of size $n = 1000$ is generated from these simulations. We then apply our MLE routine to this dataset as if all cases in the data share the same risk aversion parameter. Row 5.8 in Table 11 demonstrates that our baseline results (see row 3.1) are not sensitive to allowing for this type of preference heterogeneity – we recover the data-generating MH regime as best-fitting and we are able to distinguish it from all alternatives at the 1% confidence level using the Vuong test. Here, the parameter estimates (especially σ and θ) are distorted relative to the data-generating baseline – we have, for the MH regime, γ_{me}=.098, σ=.70, θ=9.9, ρ=.21, μ_w=.49 and γ_w= .35.

7 Into the MLE 'Black box'

7.1 Comparing actual and simulated data

In this section we use simulated data from the model at the MLE parameter estimates to help assess some of the dimensions in which the different models of financial constraints we build fail or succeed in matching the Thai data. The idea is to give a better idea why the omnibus MLE approach picks one regime in favor of another in terms of likelihood to the data, as well as to assess the fit of the highest-likelihood regime with data outside of the sub-sample on which it was estimated.

Thai vs. simulated data – assets' persistence

The different financial regimes put endogenous constraints on the ability of firms to adjust assets or, in other words, endogenize the degree of persistence of assets/capital k. For example, the FI regime stipulates that an agent, facing no financial constraints, could immediately adjust to the first-best capital level, k^* no matter what his initial k is. Such adjustment is subject to incentive compatibility constraints in the MH regime or self-enforcement constraints in the LC regime and subject to even more stringent borrowing constraints (e.g., zero borrowing under savings only and autarky) in the exogenously incomplete markets regimes. A salient feature of the Thai rural data is that investment events are infrequent (Samphantharak and Townsend, 2009). Likewise, as is evident from Table 3, capital is very persistent — the median yearly investment i, computed from the k data as $i_t = k_{t+1} - (1-\delta)k_t$, is close to zero (20 Baht). The persistence in assets favors the S (and sometimes B) regimes overall. It may also be the reason why in our robustness runs with quadratic adjustment costs the likelihood of the FI regime improves with (c,q,i,k) data.

Figure 3 helps illustrate these observations. We plot the fractions of all possible k_t to k_{t+1} transitions between any pair of points in the assets grid K over the whole panel 1999-2005. Remember the grid points correspond to the 10th, 30th, 50th, 70th and 90th percentile of k in the data. We see that in the Thai Monthly (rural) data (the top left panel) basically all transitions are on the main diagonal or the diagonals immediately next to it indicating very high persistence in assets. In contrast, in the urban data (the top right panel) there is still persistence (the main diagonal) but there are also many more households transitioning across the capital levels each year. The bottom two panels of Figure 3 plot the same transition fractions but for the simulated data at the MLE estimates (with the rural data) for the best-fitting S and MH regimes from our baseline runs. We see that the S regime manages to come the closest to the rural data in terms of the diagonal pattern (although not perfectly). The moral hazard (MH) model predicts too many off-diagonal transitions unlike what we see in the rural data but closer to what the urban data looks like. We view this as supportive, independent evidence for our results with rural data (where S wins) and urban data (where MH wins with *cqik* data and comes much closer in terms of likelihood to S with *kiq* data compared to rural) in Tables 5 and 7 respectively.

Thai vs. simulated data – time paths

Figure 4 explores how well the best fitting regime in terms of likelihood matches the paths of the mean and standard deviation (in model units) of the data variables consumption, c, business assets, k and income, q over the entire panel 1999-2005, far more data than we used in any MLE estimation. To match the data time span, we use the MLE estimates from row 5.1 in Table 5 where we used the 1999 k distribution and the 2004-05 data for c, q and i. Figure 4 shows that the best-fitting regime (here, B) tracks extremely well the time paths of the means of all three variables over the complete sample period. In terms of standard deviations, the model traces relatively well that for consumption but understates the variance of output in the Thai data and, even more significantly, the variance of assets after the initial period. The reason is the very skewed k and q distributions in the data, with few extremely large observations that are lumped on the highest grid point of the K and Q grids. When we plot the standard deviations excluding observations with average assets above the 90th percentile in the data (there are 54 such observations or 10% of the sample – see the right panels of Figure 4), the dashed lines – the standard deviations of k and q in the model come within 0.1 model units of those in the data.

Thai vs. simulated data – alternative measure of fit

We use the MLE estimates with 1999/00 (c,q,i,k) data to simulate data from each model as explained in Section 6.2. We then compute a set of 22 summary statistics or 'moments' (mean, median, standard deviation, skewness for each of the four variables c, k, i, q plus the six bilateral correlations between them) for each of the model regimes and the same statistics (in model units) for the Thai data panel years used in the MLE estimation (here, 1999 data for c, q and k and 1999 and 2000 k data used to compute investment i). Of course the MLE is not matching moments per se. Nevertheless we compute a 'goodness-of-fit' measure between the actual and simulated data as $D^m = \sum_{j=1}^{\#s} \frac{(s_j^{data} - s_j^m)^2}{|s_j^{data}|}$ where s_j^m, $j = 1, ..., 22$ denotes each of the computed moments in model m and s_j^{data} is the corresponding value in the 1999-00 Thai data. The following Table 12 (in the text) reports the value of the measure D^m for each of the six models (smaller values indicate better fit):

Table 12 – Thai vs. simulated data – mean-squared criterion

model, $m =$	**MH**	**FI**	**B**	**S**	**A**	**LC**
criterion value (rural data), $D^m =$	321.1	395.4	38.5	20.8	28.1	6520
criterion value (urban data), $D^m =$	36.8	32.0	36.4	35.3	35.4	236.7

We see that the S regime achieves the lowest D^m measure (best fit) while the LC regime has the highest measure (worst fit). This is consistent with our baseline results from Table 4 even though, to repeat, the likelihood approach uses a completely different criterion of fit. We also did the same exercise with the urban data (Table 12, third row). In contrast to the rural data, we found that a less constrained regime (the full information regime) attains the lowest criterion value (32.0) followed by the other regimes except LC all bunched in the 35-36 range.

Thai vs. simulated data – financial net worth

Figure 5 compares the implications of the best fitting regime with 1999/00 *cqik* data (the S regime, see row 3.1 in Table 5) for the time path of savings in the model and the data. Remember that in the estimation we assumed that initial savings b are distributed normally with mean and variance which are estimated, but not using the actual savings data (to maintain symmetry with the MH, FI, LC regimes). The figure compares the median and standard deviation of savings, as computed from the S model with a measure of 'financial net worth' from the data (see Pawasutipaisit et al., 2011).[25] To plot the figure and since the data is very skewed again, we exclude 'outliers' with more than three model units of savings or one model unit of debt (remember, one model unit equals the 90th percentile of business assets in the data). Figure 5 shows that the model is able to match the out-of-sample financial net worth data relatively well.

Thai vs. simulated data – return on assets

Figure 6 plots realized gross 'return on assets' (ROA), defined as income per unit of productive assets, $\frac{q}{k}$ in the Thai data, rural and urban, and compares it with the corresponding simulated values from the S and MH models computed at the MLE estimates from 1999/00 *cqik* data (Table 4, section 3). We compute the gross ROA for each year in the panel and plot the average for each household over all years against the household's average business assets holdings over that time period. We see that, for the rural data the S model which we found best-fitting in the MLE (Table 5, row 3.1) fits the general pattern (convex and downward-sloping) better than the MH model which exhibits a lot of 'bunching' in the relatively low k, low ROA range. The urban data seems to share visual features with both the MH and S panels – the bunching at low k and the hyperbola shape.

7.2 Euler equations GMM estimation

In this section we report results from two robustness estimation runs that use the Euler equations approach. These results supplement the baseline MLE results by using a different methodology to assess the fit of the models of dynamic financial constraints with the Thai data.

7.2.1 Consumption Euler equations

Following Ligon (1998) we test a moral hazard vs. 'permanent income' (borrowing and lending in a risk-free asset) models based solely on their implications for the path of consumption over time. The 'permanent income hypothesis' (PIH) (standard non-contingent debt model) implies the Euler equation,

$$u'(c_{it}) = \beta R E_t(u'(c_{it+1}))$$

[25]We did not use these data to initialize the b distribution in the estimation routine for two reasons: (i) we want to keep the B and S regimes on even ground with the MH and FI regimes where the state variable w distribution is unknown and (ii) these data only became available to us very recently, after all MLE runs were already completed.

that we estimate using our panel data on consumption, $\{c_{it}\}$, $i = 1, ...N$, $t = 1, ..T$. Suppose u is CRRA, with coefficient γ , that is, $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ and $u'(c) = c^{-\gamma}$ and let also $\beta R = 1$. Denoting $\eta_{i,t} \equiv \frac{c_{i,t+1}}{c_{i,t}}$ for $i = 1, ..N$ and $t = 1, ..T - 1$ and $h(\eta_{i,t}, b) \equiv \eta_{i,t}^b - 1$ where $b = -\gamma$ (minus the coefficient of relative risk aversion) we have the moment conditions,

$$E_t h(\eta_{i,t}, b) = 0$$

On the other hand, Rogerson (1985) or Ligon (1998) derive the corresponding 'inverse' consumption Euler equation for a repeated moral hazard model as:

$$\frac{1}{u'(c_{it})} = \frac{1}{\beta R} E_t(\frac{1}{u'(c_{it+1})})$$

Again, under CRRA and $\beta R = 1$, equation (2) can be written as:

$$E_t h(\eta_{i,t}, b) = 0 \tag{17}$$

where here $b = \gamma$ (the coefficient of relative risk aversion). As proposed by Ligon (1998), conditions (17) can be used in a GMM estimation to: (i) estimate the parameter b and (ii) use the estimate of b from step (i) to infer which model (PIH vs. private information) is likely holding in the data. Basically, assuming households are risk-averse, a positive estimate for b would indicate that private information is consistent with the data, while if the b estimate is negative, the PIH model is consistent with the data. A version of (17), $E(h(\eta_{i,t}, b).\zeta_{i,t}) = 0$ using variables $\zeta_{i,t}$ that are in the information set of household i at time t as instruments can be also estimated (see Ligon, 1998 for details).

Table 13 reports the results from the GMM estimation described above with the Thai rural data. The estimate of b is statistically significantly negative in all runs, with different sets of instruments which include pre-determined income and assets data. This result supports the PIH (non-contingent debt/savings) model, such as our B and S regimes, as opposed to the moral hazard model. As such it is not inconsistent with our baseline findings from Table 5. Note, however, that this method allows us to identify only certain parameters (the coefficient of risk aversion).[26] In addition, we are not able to distinguish between the rest of the regimes.

We also ran the same GMM test on the urban data. Unfortunately, the results are not as well-behaved as with the rural data. The run without instruments produces b=-0.22, that is (weaker) evidence in PIH favor, but with instruments included, income alone and income and capital we obtain b=8.52 and b=4.73 respectively which is evidence in favor in moral hazard (the last run from Table 13 with three instruments produced a singular matrix so we cannot report b).

[26] We also ran a version of the Ligon (1998) test where one can identify the coefficient of risk aversion and the product βR. The results (available on request) again lend support to the B/S models as opposed to the MH model.

7.2.2 Investment Euler equations

We follow Bond and Meghir (1994), we test a model of no financial constraints and quadratic adjustment costs vs. the alternative of financial constraints. Specifically, we estimate the following equation using GMM methods from Arellano and Bond (1991),

$$\left(\frac{i}{k}\right)_{jt} = \beta_1 \left(\frac{i}{k}\right)_{jt-1} + \beta_2 \left(\frac{i}{k}\right)^2_{jt-1} + \beta_3 \left(\frac{q}{k}\right)_{jt-1} + d_t + \eta_j + \varepsilon_{jt}$$

where j denotes household, t is time, and i, k, q are investment, capital and income (cash flow) respectively, as before. Bond and Meghir (1994) show that under the null of no financial constraint we must have $\beta_1 \geq 1$, $\beta_2 \leq -1$ and $\beta_3 < 0$. The focus in this literature (not without controversy, e.g. see Kaplan and Zingales, 1997) has been on the cash flow coefficient β_3. A positive β_3 estimate implying that investment, i is 'sensitive' to fluctuations in cash flow, q has been interpreted as indicating the presence of financial constraints.

Table 14 contains the results from the above estimation using the *xtabond2* function in Stata. We obtain an estimate for β_3 which is statistically insignificant from zero instead of negative. The estimates of the coefficients β_1 and β_2 also do not fall into the ranges implied by theory under the null. This indicates that we can reject the null of no financial constraints. Compared to our MLE results with *kiq* data when allowing for adjustment costs (Table 9, line 3.2) where S and A with adjustment costs are tied for best fit we reach a similar conclusion which we view as further supporting evidence for our MLE findings in favor of the exogenously incomplete financial regimes when using the investment, income and assets data. However, unlike our structural MLE approach, the investment GMM method does not allow us to distinguish across the multiple possibilities of sources for financial constraints (as modeled in the A through MH regimes).

8 Conclusions

We formulate and solve numerically a wide range of models of dynamic financial constraints with exogenous or endogenous contract structure that allow for moral hazard, limited commitment and unobservable output, capital and investment. We develop methods based on mechanism design theory and linear programming and used them to structurally estimate, compare and statistically test between the different financial regimes. The compared regimes differ significantly with respect to their implications for investment and consumption smoothing in the cross-section and dynamics. Combined consumption and investment data were found particularly useful in pinning down the financial regime. Our methods can handle unobserved heterogeneity, grid approximations, transitional dynamics, and reasonable measurement error.

We have established that our methods work on actual data from Thailand. We echo previous work which finds that full risk sharing is rejected, but not by much, and indeed find that the moral hazard regime is not inconsistent with the rural income and consumption data and with the joint business and consumption data in our urban sample. We also recover a more sophisticated contract theoretic regime (moral hazard constrained credit) if we restrict attention to family or gift/loan networks data, confirming related work by Chiappori et al. (2009) and Kinnan and Townsend (2009).

In terms of investment, we confirm previous work which finds that investment is not smooth and may be sensitive to cash flow and, indeed, find that more constrained regimes such as saving only and borrowing and lending characterize best the investment and income data (both rural and urban), as well as the combined consumption, income and investment data in the rural sample. The reasons for these findings are likely due to the infrequent nature of investment in the Thai data (especially the rural sample) and the relatively large size of investment compared to capital when investment takes place. On the other hand, the financial regimes we study postulate endogenous constraints on the ability of firms to adjust assets, embedding the degree of assets persistence. The feature of the Thai data that capital is persistent thus favors the S (or B) regimes where assets adjustment is subject to more stringent constraints than in MH or FI. Evidently we have learned something from our approach, beginning to distinguish, in a sense, capital adjustment costs from financial constraints.

The results can be put in perspective relative to our previous findings in Paulson, Townsend and Karaivanov (2006) where we estimated a one-period model of binary occupational choice between starting a business and subsistence farming as a function of ex-ante wealth. We found moral hazard (rather than limited liability) to be the predominant source of financial constraints for rural Thai entrepreneurs, but the borrowing and saving only regimes we study here were not tested. In contrast, in the current paper we not only introduce full-blown dynamic mechanism design models but also significantly expand our previous work and estimate using cross-sections or panels of consumption, investment, assets and income, separately and jointly. On the other hand, Karaivanov (2012) finds that, in an occupational choice setting similar to Paulson et al., one cannot distinguish statistically between a model of moral hazard vs. a model of borrowing with default similar to what we find in the (c, q) cross-section specifications in this paper. The reason for the differences in the results are the different data on which we fit, which here, using joint distribution of data variables in the MLE is much richer than binary occupational choice used in our previous work. In this earlier work we model and estimate models of an one-off constraint on business start-ups as opposed the fully dynamic constraints on consumption and investment smoothing here.

One important finding from our robustness runs with data simulated from one of the models is that using combined consumption and investment data we can readily distinguish exogenously incomplete financial regimes from endogenously incomplete ones, where the latter are solutions to mechanism design problems with unobserved actions and state variables. As the literature we surveyed in the introduction typically takes one route or the other, we believe this ability to distinguish will prove helpful in future research and the applications of others. We are also able to distinguish within these regime groups, though this depends on measurement error, the variables in the available data set, and whether or not we have more than a single cross-section of data. Of course, we do not claim that we have covered all possible models of financial constraints, only six common prototypes. Obvious inclusions left for future work are model with observed effort but unobserved ability or productivity or models with aggregate shocks. Other natural future steps include allowing for distinctions across different technologies (fish, shrimp, livestock, business, etc.) and aggregate shocks (shrimp disease, rainfall, etc.). We would also like to return to the issue of entrepreneurial talent, as in our earlier work (Paulson et al., 2006) and allow for heterogeneity in project returns. Related work (Pawasutipaisit and Townsend, 2008) shows that ROA varies considerably across households and is persistent. On the other hand, such data summaries have trouble finding consistent patterns with respect to finance, suggesting the data be viewed through the lens of revised models.

We are still somewhat limited on the computational side, though we are encouraged with the recent advances we have made. We had difficulty estimating the moral hazard regime with unobserved capital and investment. In an on-going collaboration with computer scientists, we have been exploring the use of parallel processing to speed up our codes and allow more complexity. What we have done thus far is, for want of better terminology, brute force. There would be further gains from more streamlined programs and more efficient search, i.e., where to refine the grids, when to use non-linear or mixed methods, the use of nested pseudo-likelihood methods, and so on.

We have our eyes on other economies as well, in part because we get more entry and exit from business in other countries, and in part because we need large sample sizes for our methods to work. Unfortunately, we do not typically find both consumption and investment data, which is why we chose the Thai data to begin with. Work in progress (Karaivanov, Ruano, Saurina and Townsend, 2009) with non-financial firms data from Spain shows evidence that the number of firms' bank relationships matters for whether they exhibit excess cash flow sensitivity of investment. We use the methods described in the current paper to evaluate which of four financial regimes (autarky, non-contingent debt, moral hazard and complete markets) best characterizes the degree of financial constraints for unbanked, single-banked and multiple-banked firms. Our methods allow in principle for transitions across financial regimes which is another extension we plan.

References

[1] Abraham, A. and N. Pavoni (2008), "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending: A Recursive Formulation", *Review of Economic Dynamics* (forthcoming).

[2] Abel, A. and O. Blanchard (1983), "An Intertemporal Model of Savings and Investment", *Econometrica*, 51(3), pp. 675-92.

[3] Abreu, D., D. Pearce and E. Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring", *Econometrica* 58(5), pp. 1041-63.

[4] Aiyagari, S.R. (1994), "Uninsured Idiosyncratic Risk and Aggregate Saving", *Quarterly Journal of Economics*, 109(3), pp. 659-84

[5] Albuquerque, R. and H. Hopenhayn (2004), "Optimal Lending Contracts and Firm Dynamics," *Review of Economic Studies*, 71(2), pp. 285-315.

[6] Alem, M. and R. Townsend (2009), "An Evaluation of Safety Nets and Financial Institutions in Crisis and Growth", working paper, MIT.

[7] Angeletos, G. and L. Calvet (2007), "Idiosyncratic Production Risk, Growth, and the Business Cycle", working paper, MIT.

[8] Attanasio, O. and N. Pavoni (2008), "Risk Sharing in Private Information Models with Asset Accumulation", working paper, UCL.

[9] Berger, A. and G. Udell (2002), "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure", *Economic Journal*, 112, pp. F32-F53.

[10] Bitler, M., T. Moskowitz and A. Vissing-Jorgensen (2005), "Testing Agency Theory with Entrepreneur Effort and Wealth", *Journal of Finance*, 60(2), pp. 539-76.

[11] Bond, S. and C. Meghir (1994), "Dynamic investment models and the firms' financial policy", *Review of Economic Studies*, 61(2), pp. 197-222.

[12] Buera, F. and Y. Shin (2007), "Financial Frictions and the Persistence of History: A Quantitative Exploration", working paper, Northwestern.

[13] Cagetti, M. and M. De Nardi (2006), "Entrepreneurship, Frictions, and Wealth", *Journal of Political Economy*, 114(5), pp. 835-70

[14] Chiappori, P., S. Schulhofer-Wohl, K. Samphantharak, and R. Townsend (2008), "Heterogeneity and Risk Sharing in Thai Villages", working paper, University of Chicago.

[15] Clementi, G. and H. Hopenhayn (2006), "A Theory of Financing Constraints and Firm Dynamics", *Quarterly Journal of Economics*, 121(1), pp. 229-65.

[16] Cooley, T. and V. Quadrini (2001), "Financial Markets and Firm Dynamics", *American Economic Review*, 91(5), pp. 1286-1310.

[17] Covas, F. (2006), "Uninsured idiosyncratic production risk with borrowing constraints", *Journal of Economic Dynamics and Control*, 30(11), pp. 2167-90.

[18] Deaton, A. and G. Laroque (1996), "Competitive Storage and Commodity Price Dynamics", *Journal of Political Economy*, 104(5), pp. 896-923.

[19] Doepke, M. and R. Townsend (2006), "Dynamic Mechanism Design with Hidden Income and Hidden Actions", *Journal of Economic Theory,* 126(1), pp. 235-85.

[20] Dubois, P., B. Julien and T. Magnac (2008), "Formal and Informal Risk Sharing in LDCs: Theory and Empirical Evidence", *Econometrica*, 76(4), pp. 679-25.

[21] Ericson, R. and A. Pakes, (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", *Review of Economic Studies*, 62(1), pp. 53-82.

[22] Fazzari, S., R. Hubbard and B. Petersen (1988), "Financing Constraints and Corporate Investment", *Brooking Papers on Economic Activity*, 1, pp. 141-206.

[23] Fernandes, A. and C. Phelan (2000), "A Recursive Formulation for Repeated Agency with History Dependence", *Journal of Economic Theory*, 91, pp. 223-247.

[24] Geanakoplos, J. and H. Polemarchakis (1986), "Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When Markets are Incomplete", in: W.P. Heller, R.M. Ross, D.A. Starrett (eds.): *Uncertainty Information and Communication. Essays in Honor of Kenneth Arrow Vol. 3*, Cambridge.

[25] Heaton, J. and D. Lucas, (2000), "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk", *Journal of Finance*, 55(3), pp. 1163-98.

[26] Honore, B. and E. Tamer (2006), "Bounds on Parameters in Panel Dynamic Discrete Choice Models", *Econometrica*, 74(3), pp. 611-30.

[27] Hopenhayn, H. (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(2), pp. 1127-50.

[28] Jappelli, T. and L. Pistaferri (2006), "Intertemporal Choice and Consumption Mobility", *Journal of the European Economic Association*, 4(1), pp. 75-115.

[29] Jewitt, I. (1988), "Justifying the First-Order Approach to Principal-Agent Problems,", *Econometrica*, 56(5), pp. 1177-90.

[30] Judd, K. (1998), *Numerical Methods in Economics*, MIT Press, Cambridge, MA.

[31] Kaplan and Zingales (2000), "Investment and Cash Flow Sensitivities are not Valid Measures of Financing Constraints", *Quarterly Journal of Economics*, 115(2), pp. 707-12.

[32] Karaivanov, A. (2012), "Financial Constraints and Occupational Choice in Thai Villages", *Journal of Development Economics*, forthcoming.

[33] Karaivanov, A., S. Ruano, J. Saurina and R. Townsend (2009), "Formal Financial Access for Older Firms; Does it Matter for Firms' Investment?", working paper.

[34] Kinnan, C. and R. Townsend (2009), "Kinship Networks, Financial Access and Consumption Smoothing", working paper, MIT

[35] Kocherlakota, N. (2004), "Figuring Out the Impact of Hidden Savings on Optimal Unemployment Insurance", *Review of Economic Dynamics* 7(3), pp. 541-554.

[36] Ligon, E. (1998), "Risk Sharing and Information in Village Economies", *Review of Economic Studies* 66(4), pp. 847-64.

[37] Ligon, E., J. Thomas and T. Worrall, (2002) "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies", *Review of Economic Studies*, 69(1), pp. 209-44.

[38] Meh, C. and V. Quadrini (2006), "Endogenous Incomplete Markets With Investment Risks", *Journal of Economic Dynamics and Control*, 30(11), pp. 2143-65.

[39] Myers, S. and N. Majuf (1984), "Corporate financing and investment decisions when firms have information that investors do not have", *Journal of Financial Economics*, 13(2), pp. 41-56.

[40] Paulson, A., R. Townsend and A. Karaivanov (2006), "Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship", *Journal of Political Economy*, 114(1), pp. 100-144.

[41] Pawasutipaisit, A. and R. Townsend (2008), "Enterprise and the Growth of Wealth", working paper, MIT.

[42] Phelan, C. (1994), "Incentives, Insurance and the Variability of Consumption and Leisure", *Journal of Economic Dynamics and Control*, 18, pp. 581-99.

[43] Phelan, C. and R. Townsend (1991), "Computing Multi-Period, Information-Constrained Equilibria", *Review of Economic Studies*, 58, pp. 853-881.

[44] Prescott, E.C. and R. Townsend (1984), "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard," *Econometrica*, 52, pp. 21–45.

[45] Prescott, E. S. (2003), "Communication in Models with Private Information: Theory and Computation", *Geneva Papers on Risk and Insurance Theory* 28(2), pp. 105-30.

[46] Rivers, D. and Q. Vuong, (2002), "Model Selection Tests for Nonlinear Dynamic Models", *The Econometrics Journal*, 5, pp. 1-39.

[47] Rogerson, W. (1985), "The First Order Approach to Principal-Agent Problems", *Econometrica* 53, pp. 1357-68.

[48] Rossi-Hansberg, E. and M. Wright (2007), "Establishment Size Dynamics in the Aggregate Economy", *American Economic Review*, 97(5), pp. 1639-66.

[49] Samphantharak, K. and R. Townsend (2009), *Households as Corporate Firms: An Analysis of Household Finance Using Integrated Household Surveys and Corporate Financial Accounting*, Cambridge University Press.

[50] Schmid, L. (2008), "A Quantitative Dynamic Agency Model of Financing Constraints", working paper, Duke University.

[51] Spear, S. and S. Srivastava (1987), "On Repeated Moral Hazard with Discounting", *Review of Economic Studies,* 53, pp. 599-617.

[52] Townsend, R. (1982), "Optimal Multi-period Contracts and the Gain from Enduring Relationships Under Private Information", *Journal of Political Economy* 61, pp. 166-86.

[53] Townsend, R., A. Paulson and T. Lee (1997), *Townsend Thai Project Surveys*, http://cier.uchicago.edu

[54] Vuong, Q. (1989), "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses." *Econometrica*, 57(2), pp. 307-333.

[55] Werning, I. (2001), "Repeated Moral Hazard with Unmonitored Wealth: A Recursive First Order Approach", working paper, MIT.

[56] Zame, W. (1993), "Efficiency and the Role of Default when Security Markets are Incomplete", *American Economic Review*, 83, pp. 1142-64.

[57] Zeldes, S. (1989), "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence", *Quarterly Journal of Economics*, 104(2), pp. 275-98.

9 Appendices

9.1 Appendix A – Moral hazard with unobserved capital and investment

Let agent's effort be unobservable as in the moral hazard (MH) regime but, suppose in addition, the principal also cannot observe the agent's current capital stock k and planned level for next period,

k'. The unobserved state k adds a dynamic adverse selection problem to the moral hazard problems arising from the two unobserved actions, z and k'.

To model this setting as a mechanism-design problem, suppose the agent sends a message about his capital level k to the principal who offers him a contract conditional on the message which consists of transfer τ, recommended effort z, investment k', and future promised utility. Because of the dynamic adverse selection problem in the state k, following Fernandes and Phelan (2000) and Doepke and Townsend (2006), instead of the scalar promise w in the MH regime, the proper state variable in the recursive representation is a *promised utility schedule*, $\mathbf{w} \equiv \{w(k_1), w(k_2), ..., w(k_{\#K})\} \in \mathbf{W}$, where k_1, k_2, etc. are the elements of the grid K.[27] The $\#K$-dimensional set $\mathbf{W}$ is endogenously determined (not all promise-assets combinations are feasible) and must be iterated upon together with the value and policy functions (Abreu, Pierce and Stacchetti, 1990).

The computational method we use to solve for the optimal contract in this unobserved investment (UI) regime requires separability in consumption and leisure, $U(c, z) = u(c) - d(z)$ (note, this was not needed for the MH, FI, or the exogenously incomplete regimes). The separability allows us to split each time period into two sub-periods and use dynamic programming within the time periods. This helps keep dimensionality in check, since the resulting sub-problems are of much lower dimension. The first sub-period includes the announcement of k by the agent, the principal's effort recommendation z, the agent's actual effort supply, and the realization of the output q. The second sub-period includes the transfer, the investment recommendation, and the agent's consumption and actual investment decisions. To tie the two sub-periods together, we introduce the extra variables, $\mathbf{w}_m$ that we call 'interim promised utility' – a representation of the agent's expected utility from the end of sub-period 1 (that is, from the middle of the period) onwards. The interim promised utility is a *schedule* (vector), $\mathbf{w}_m = \{w_m(k_1), w_m(k_2), ...\} \in \mathbf{W}_m$, similar to $\mathbf{w}$. Like $\mathbf{W}$, the set $\mathbf{W}_m$ is endogenously determined along the value function iteration.

The first sub-period problem for computing the optimal contract with an agent who has announced k and has been promised $\mathbf{w}$ is:

Program UI1

$$V(\mathbf{w}, k) = \max_{\{\pi(q,z,\mathbf{w}_m|\mathbf{w},k)\}} \sum_{Q\times Z\times \mathbf{W}_m} \pi(q, z, \mathbf{w}_m|\mathbf{w},k)[q + V_m(\mathbf{w}_m, k)] \tag{18}$$

The choice variables are the probabilities over allocations $(q, z, \mathbf{w}_m) \in Q \times Z \times \mathbf{W}_m$. The function $V_m(\mathbf{w}_m, k)$ is defined in the second sub-period problem (see Program UI2 below). The maximization in (18) is subject to the following constraints. First, the optimal contract must deliver the promised utility on the equilibrium path, $w(k)$:

$$\sum_{Q\times Z\times \mathbf{W}_m} \pi(q, z, \mathbf{w}_m|\mathbf{w},k)[-d(z) + w_m(k)] = w(k) \tag{19}$$

The utility from consumption and discounted future utility are incorporated in w_m. Second, as in the MH regime, the optimal contract must satisfy incentive compatibility in effort. That is, $\forall(\bar{z}, \hat{z}) \in Z\times Z$:

$$\sum_{Q\times \mathbf{W}_m} \pi(q, \bar{z}, \mathbf{w}_m|\mathbf{w},k)[-d(\bar{z}) + w_m(k)] \geq \sum_{Q\times \mathbf{W}_m} \pi(q, \bar{z}, \mathbf{w}_m|\mathbf{w},k)\frac{P(q|\hat{z}, k)}{P(q|\bar{z}, k)}[-d(\hat{z}) + w_m(k)] \tag{20}$$

Third, since the state k is private information, the agent needs incentives to reveal it truthfully. On top of that, the agent can presumably consider joint deviations in his announcement, k and his effort

[27]The reason why utility promises must, in general, depend on the state k is the different incentives of agents entering next period with different capital levels (see Kocherlakota, 2004 for a detailed discussion).

choice, z. To prevent such joint deviations, truth-telling must be ensured to hold regardless of whether the agent decides to follow the effort recommendation, z or considers a deviation to another effort level $\delta(z) \in Z$, where $\delta(z)$ denotes all possible mappings from recommended to actual effort, that is, from the set Z to itself. Such behavior is ruled out by imposing the following 'truth-telling' constraints, which must hold for all $\hat{k} \neq k$ and $\delta(z)$:

$$w(\hat{k}) \geq \sum_{Q \times Z \times \mathbf{W}_m} \pi(q, z, \mathbf{w}_m | \mathbf{w}, k) \frac{P(q|\delta(z), \hat{k})}{P(q|z, k)} [-d(\delta(z)) + w_m(\hat{k})] \tag{21}$$

In words, an agent who actually has $\hat{k}$ but considers announcing k triggering $\pi(.|\mathbf{w}, k)$ should find any such deviation unattractive. There are $(\#K-1)\#Z^{\#Z}$ such constraints in total. Finally, the contract must satisfy the already familiar technological consistency, adding-up, and non-negativity constraints for the probabilities $\pi(q, z, \mathbf{w}_m|\mathbf{w}, k)$.

To solve Program UI1, we first need to compute the principal's 'interim value function' $V_m(\mathbf{w}_m, k)$. The proper state variables are the schedule, $\mathbf{w}_m$ of interim utilities for each $k \in K$ and the agent's actual announcement k. Constraints will introduce truth-telling and obedience in the second-stage program. We need to ensure that, when deciding on k', the agent cannot obtain more than his interim utility, $w_m(k)$ for any announcement k.

Program UI2

$$V_m(\mathbf{w}_m, k) = \max_{\{\pi(\tau, k', \mathbf{w}'|\mathbf{w}_m, k)\}, \{v(\hat{k}, k', \tau|\mathbf{w}_m, k)\}} \sum_{T \times K \times \mathbf{W}} \pi(\tau, k', \mathbf{w}'|\mathbf{w}_m, k)[-\tau + (1/R)V(k', \mathbf{w}')] \tag{22}$$

Note that, in addition to the allocation lotteries, $\pi(\tau, k', \mathbf{w}'|\mathbf{w}_m, k)$ we introduce additional choice variables, $v(\hat{k}, k', \tau|\mathbf{w}_m, k)$ that we refer to as 'utility bounds' (see Prescott, 2003 for details). These bounds specify the maximum expected utility that an agent who is actually at $\hat{k}$ receiving transfer τ and an investment recommendation k' could obtain by reporting k and doing $\hat{k}'$. This translates into the constraint:

$$\sum_{\mathbf{W}} \pi(\tau, k', \mathbf{w}'|\mathbf{w}_m, k)[u(\tau + (1-\delta)\hat{k} - \hat{k}') + \beta w'(\hat{k}')] \leq v(\hat{k}, k', \tau|\mathbf{w}_m, k) \tag{23}$$

which must hold for all possible combinations $\tau, k', \hat{k} \neq k$, and $\hat{k}' \neq k'$. To ensure truth-telling, the interim utility $w_m(\hat{k})$ that the agent obtains in the second sub-period by reporting k when the true state is $\hat{k}$, must satisfy, for all $k, \hat{k}$:

$$\sum_{T \times K} v(\hat{k}, k', \tau|\mathbf{w}_m, k) \leq w_m(\hat{k}) \tag{24}$$

The two sets of constraints, (23) and (24) rule out any joint deviations in the report k and the action k'. Finally, by definition, the interim utility must satisfy:

$$w_m(k) = \sum_{T \times K \times \mathbf{W}} \pi(\tau, k', \mathbf{w}'|\mathbf{w}_m, k)[u(\tau + (1-\delta)k - k') + \beta w'(k')] \tag{25}$$

and the probabilities $\pi(\tau, k', \mathbf{w}'|\mathbf{w}_m, k)$ must satisfy non-negativity and adding-up.

9.2 Appendix B – Hidden output

This is a version of the model where we allow output, q to be unobservable to the financial intermediary, similarly to Townsend (1982) or Thomas and Worrall (1990). Assume effort, z is contractible so there is no problem with joint deviations. We have:

$$V(w,k) = \max_{\{\pi(\tau,q,z,k',w'|w,k)\}} \sum_{T\times Q\times Z\times K\times W} \pi(\tau,q,z,k',w'|w,k)[-\tau + (1/R)V(w',k')]$$

subject to the *promise keeping* constraint:

$$\sum_{T\times Q\times Z\times K\times W} \pi(\tau,q,z,k',w'|w,k)[U(q+\tau+(1-\delta)k-k',z)+\beta w'] = w$$

and the *truth-telling constraints* (true output is $\bar{q}$ but the agent considers announcing $\hat{q}$), $\forall\ (\bar{z},\bar{q},\hat{q} \neq \bar{q} \in Z\times Q\times Q)$:

$$\sum_{T\times K\times W} \pi(\tau,\bar{q},\bar{z},k',w'|w,k)[U(\bar{q}+\tau+(1-\delta)k-k',\bar{z})+\beta w'] \geq$$
$$\geq \sum_{T\times K\times W} \pi(\tau,\hat{q},\bar{z},k',w'|w,k)[U(\bar{q}+\tau+(1-\delta)k-k',\bar{z})+\beta w']$$

subject to the technological consistency and adding-up constraints:

$$\sum_{T\times K\times W} \pi(\tau,\bar{q},\bar{z},k',w'|w,k) = P(\bar{q}|\bar{z},k) \sum_{T\times Q\times K\times W} \pi(\tau,q,\bar{z},k',w'|w,k) \text{ for all } (\bar{q},\bar{z}) \in Q\times Z \text{ and}$$
$$\sum_{T\times Q\times Z\times K\times W} \pi(\tau,q,z,k',w'|w,k) = 1,$$

as well as non-negativity: $\pi(\tau,q,z,k',w'|w,k) \geq 0$ for all $(\tau,q,z,k',w') \in T\times Q\times Z\times K\times W$.

9.3 Appendix C – Computing joint distributions of model variables

Our linear programming solution method allows us to easily map to likelihoods and take to data the implications of the different model regimes through the policy functions, the probabilities $\pi^*(.)$ that solve the dynamic programs in section 2. We first construct the state transition matrix for each regime. Denote by $s \in S$ the current state – k in autarky, (b,k) in S/B, or (w,k) in the MH/FI/LC regimes. The transition probability of going from any current state s to any next-period state s' is computed from the optimal policy $\pi^*()$, integrating out all non-state variables. For example, for the MH regime we have:

$$Prob(w',k'|w,k) = \sum_{C\times Q\times Z} \pi^*(c,q,z,k',w'|w,k)$$

where we have replaced the transfer τ by consumption c (the problem is mathematically equivalent). Putting these transition probabilities together for all $s \in S$ yields the state transition matrix $\mathbf{M}$ of dimension $\#S\times\#S$, (for example, for MH $\#S = \#K\times\#W$), with elements m_{ij}, $i,j = 1,...,\#S$ corresponding to the transition probabilities of going from state s_i to state s_j in S.

The matrix $\mathbf{M}$ completely characterizes the dynamics of the model. For example, we can use $\mathbf{M}$ to compute the cross-sectional distribution over states at any time t, $\mathbf{H}_t(s) \equiv (h_t^1,..,h_t^{\#S})$, starting from an arbitrary given initial state distribution, $\mathbf{H}_0(s)$ as:

$$\mathbf{H}_t(s) = (\mathbf{M}')^t \mathbf{H}_0(s) \tag{26}$$

In our empirical application we take $H_0(s)$ from the data. In practice some elements of the state s may be unobservable to the researcher, e.g. here, the state variable w in the MH, LC and FI regimes. We assume that the unobserved state is drawn from some known distribution, the parameters of which we estimate.

We use the state probability distribution (26) in conjunction with the policy functions $\pi^*(.)$ to compute cross-sectional probability distributions $F_t(x)$ for any vector of model variables x (which could include k, k', z, τ, q, c, etc.), at any time period. For example, in the MH regime, the time-t joint cross-sectional distribution of consumption over the grid C with elements c_l, $l = 1, .., \#C$ and income q over the grid Q with elements q_h, $h = 1, .., \#Q$ is:

$$F_t(c_l, q_h) \equiv \text{Prob}_t(c = c_l,\ q = q_h | \mathbf{H}_0) = \sum_{j=1..\#S} h_t^j \sum_{Z \times K \times W} \pi_t^*(c = c_l, q = q_h, z, k', w' | s^j)$$

We also use the time-t distribution over states $\mathbf{H}_t(s)$ and the transition matrix $\mathbf{M}$ to compute the transition probabilities, $\mathbf{P}_t(x, x')$ for any model variable x, at any time period, t. The transition and the cross-sectional probabilities are then easily combined to construct joint probability distributions encompassing several periods at a time as in a panel.

9.4 Appendix D – Simulating data from the models

To simulate data from the moral hazard (MH) regime we fixed the parameter values as follows: risk aversion, $\sigma = .5$, effort curvature, $\theta = 2$ and the production function parameter $\rho = 0$ (corresponding to Cobb-Douglas form). These parameters are representative, chosen from a large set of runs we did and generate well-behaved interior solutions for the baseline grids chosen (we use a five-point k grid on $[0, 1]$, five-point q grid on [.05,1.75]). The rest of the parameters are the same as discussed in Section 5.2.

We simulate data from the MH model at the baseline parameters, ϕ^{base} and grids described above. We take an initial distribution over the states (k, w) which has an equal number of data points for each grid point in the capital grid K and is normally distributed in w, i.e., $w \sim N(\mu_w, \gamma_w^2)$ for each $k \in K$.[28] We set the mean μ_w to be equal to the average value in the promise grid, $\frac{w_{\max}+w_{\min}}{2}$, at the baseline parameters; the standart deviation is set to 0.35 of the w grid span. We next compute the data-generating regime (MH) at the baseline parameters, ϕ^{base} given the drawn initial distribution over states (k, w) as described above and use the LP solution π^* to generate, via a Monte Carlo procedure (with fixed random numbers across regimes) draw 'data' on c, q, k and i to use in the estimation.

Consistent with the runs with actual data, the simulatuon allows for additive normally distributed measurement error in all variables. We perform all estimation runs and tests in this section for two measurement error specifications: 'low measurement error', where we set the parameter $\tilde{\gamma}_{me}$ governing the size of measurement error relative to a variable's range equal to 0.1, and 'high measurement error', with $\tilde{\gamma}_{me} = 0.2$ of the grid range.

[28] Our methods allow any other possible initializations (mixtures of normals or bivariate distributions), at the cost of additional parameters to be estimated and slower computation. In the Thai data application in the next section we use the actual initial discretized distribution of assets in the data.

Figure 1: Thai data − income, consumption, investment comovement

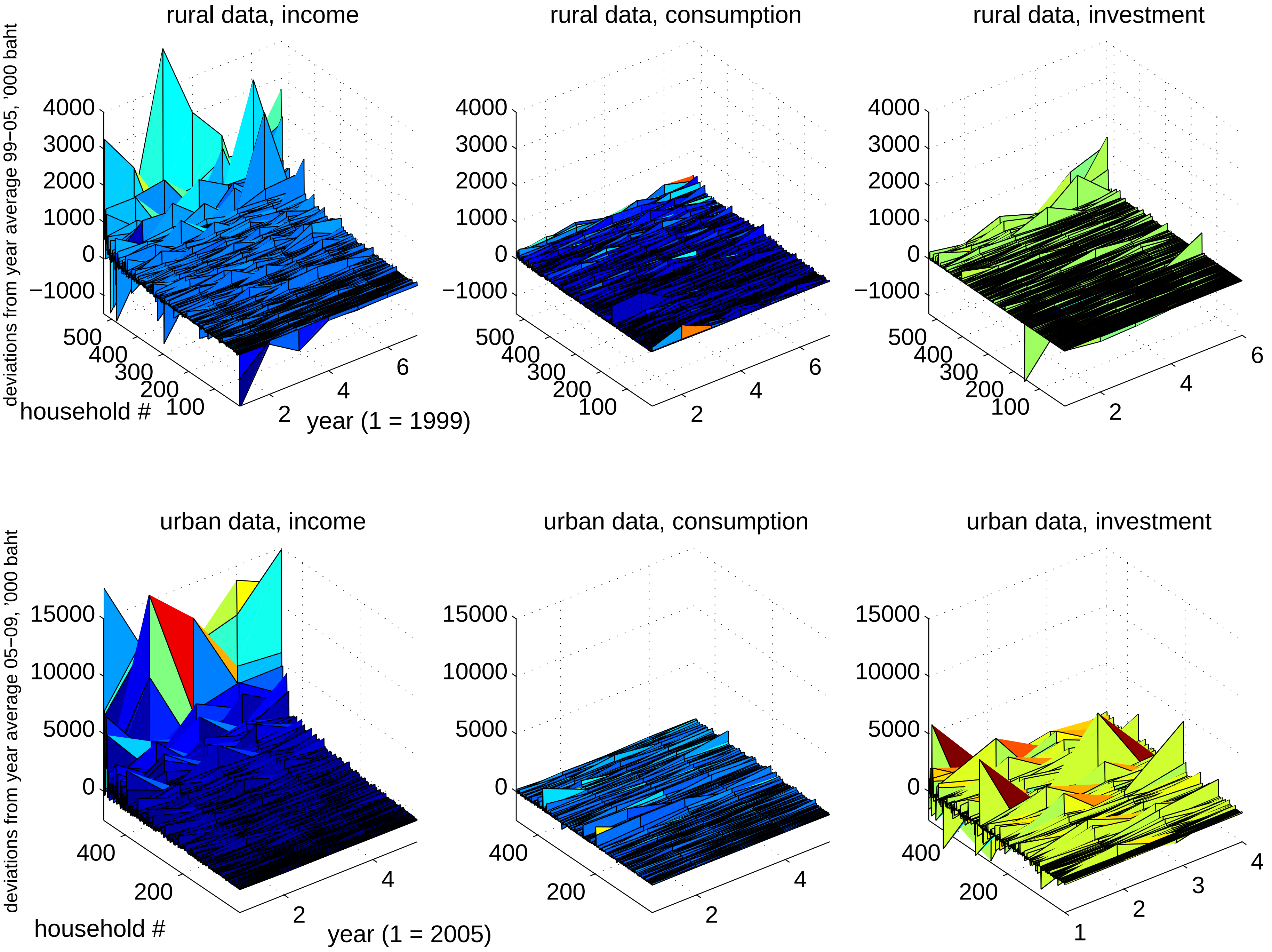


Figure 2: Thai data − income, consumption, assets changes

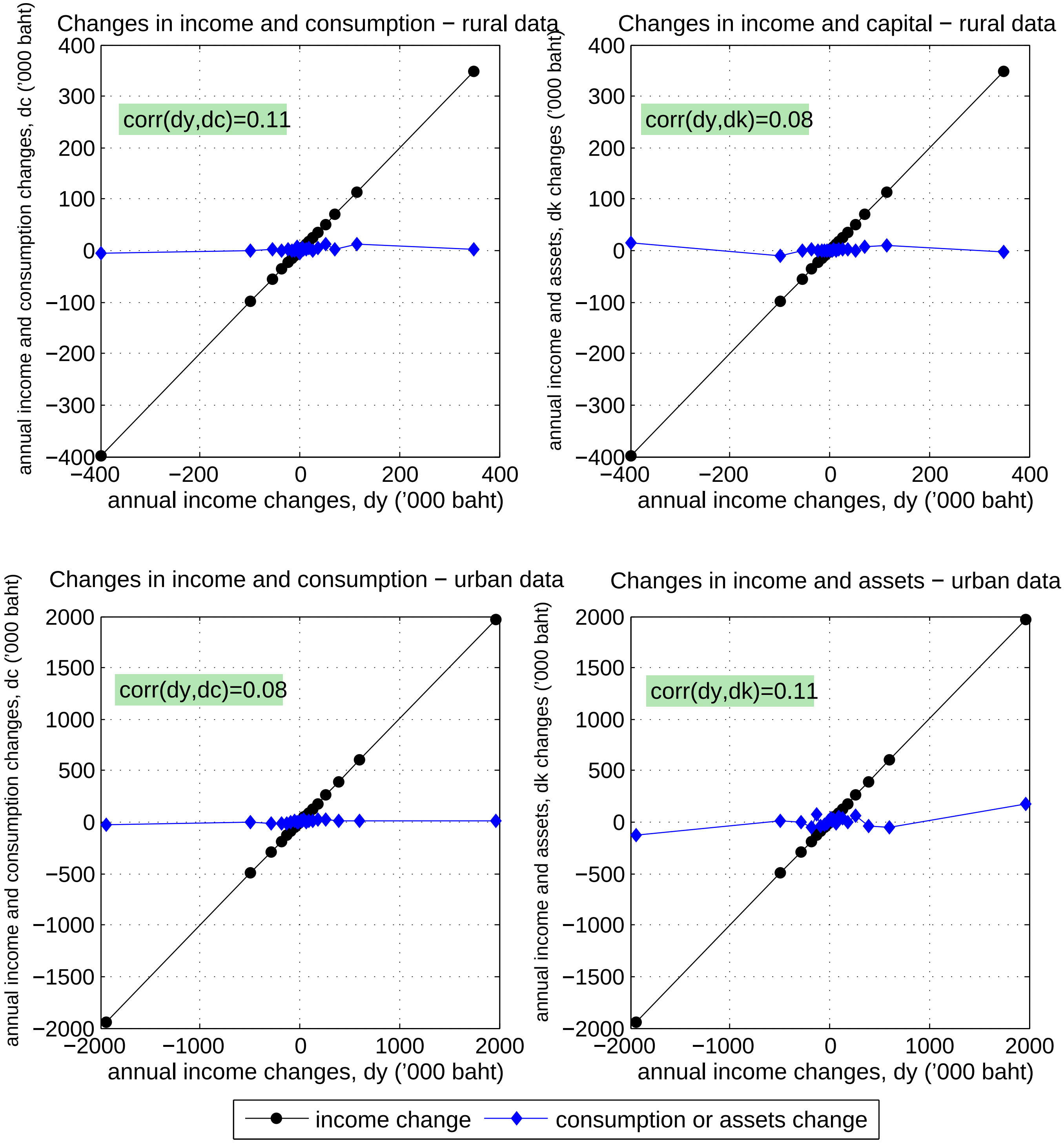
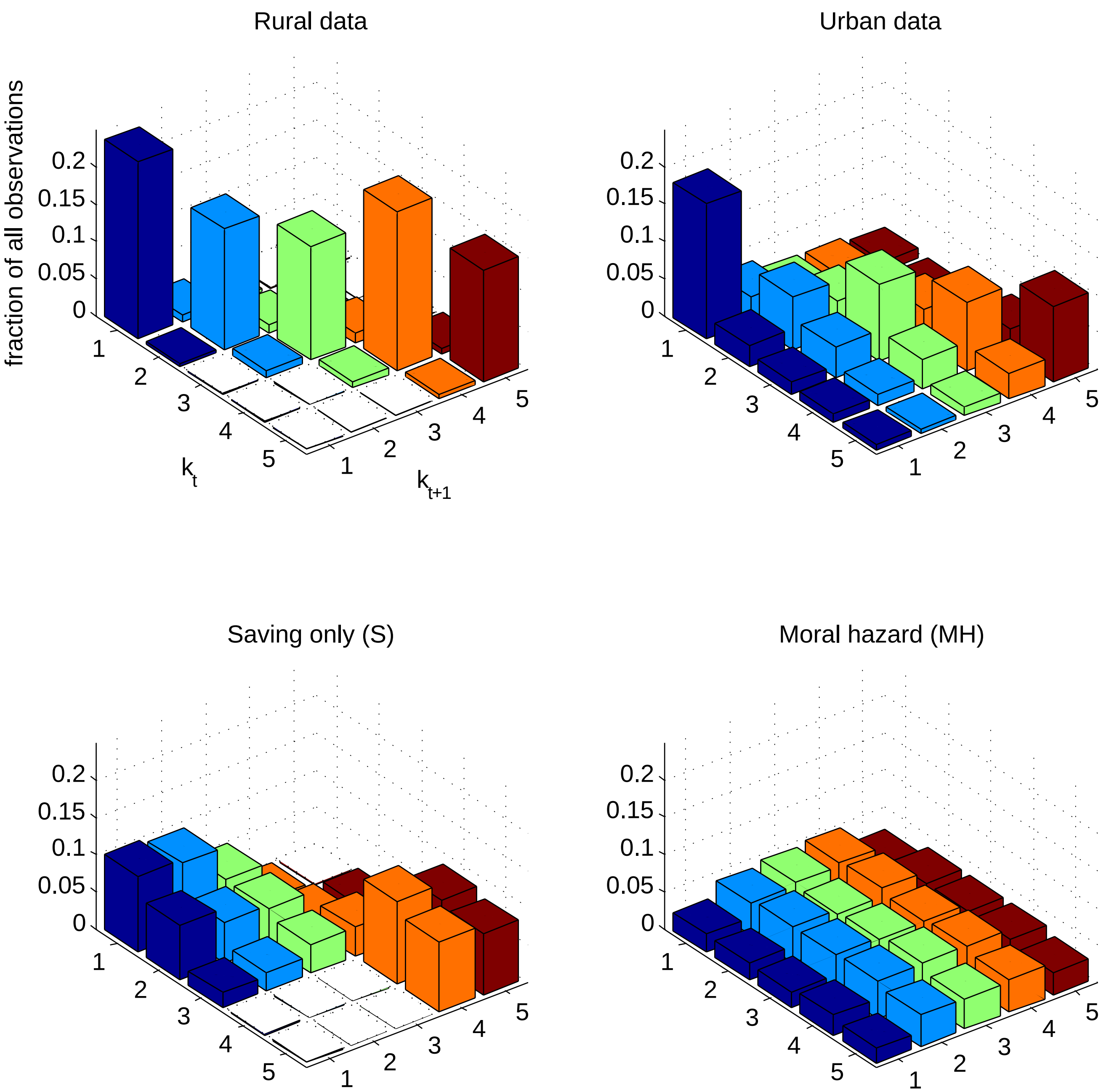


Figure 3: Thai vs. simulated data; business assets transition matrix



Note: axis labels corresponds to k percentiles; 1 is 10th, 5 is 90th; values larger than $4*10^{-3}$ plotted in color

Figure 4: Thai vs. Simulated data − Time Paths

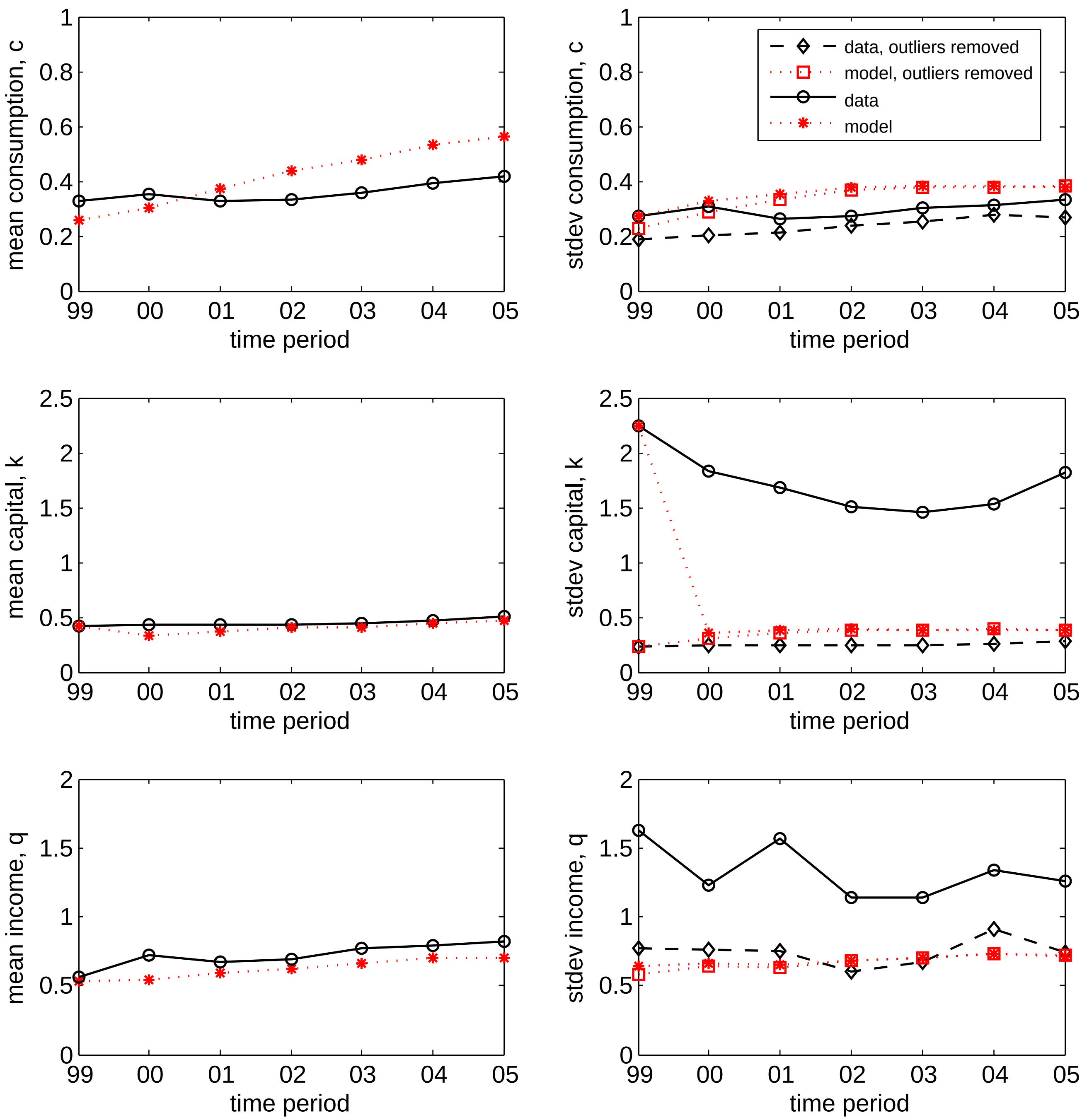


Figure 5: Thai vs. simulated data − savings

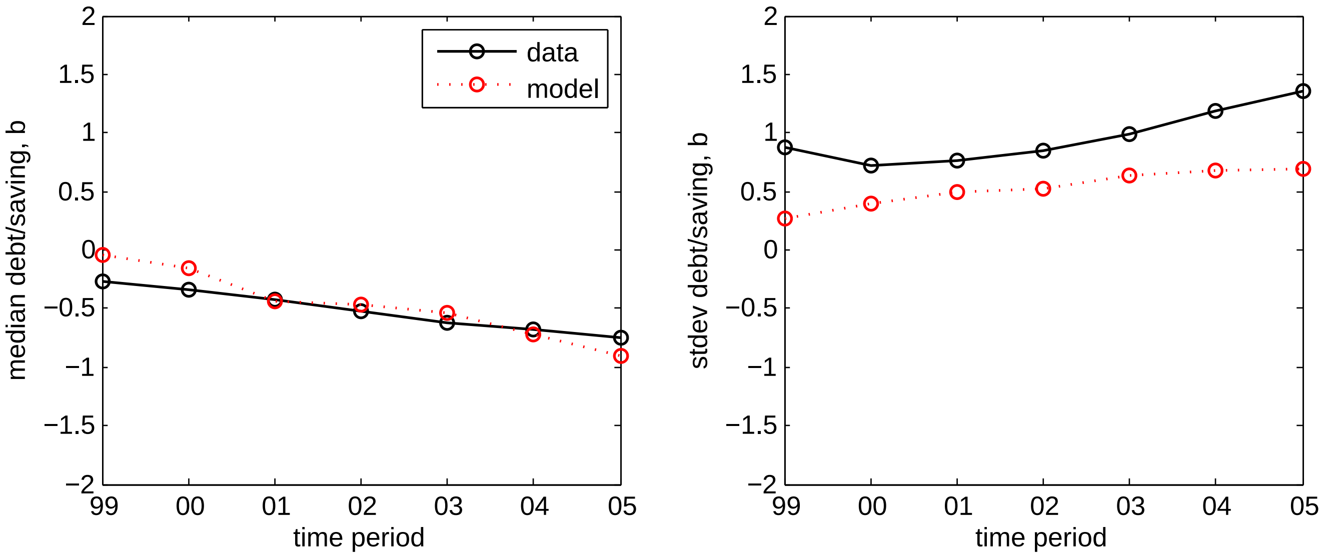


Figure 6 − Thai vs. simulated data − return on assets

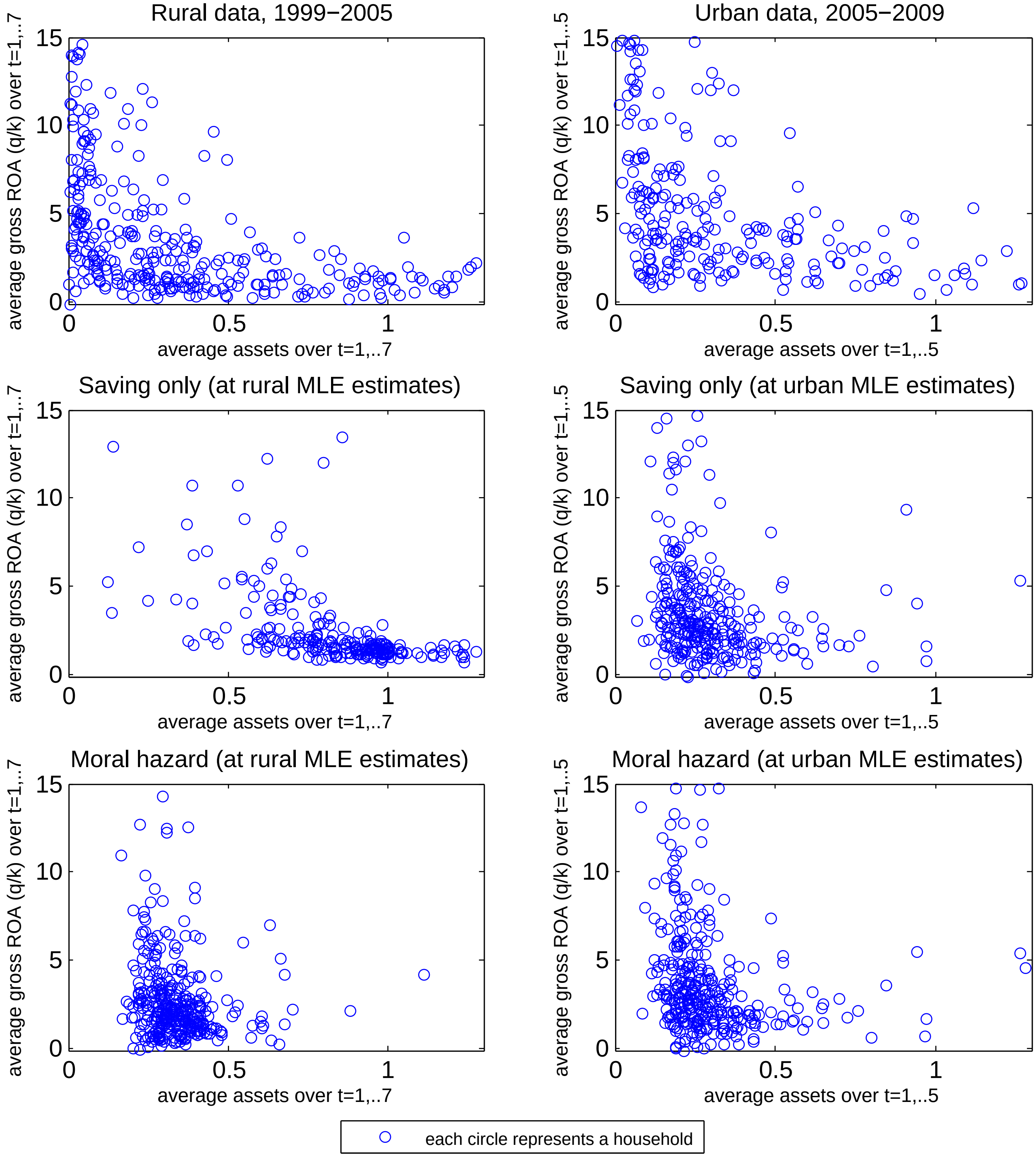


Table 1 - Problem Dimensionality

Number of: Model:	linear programs solved per iteration	variables per linear program	constraints per linear program
Autarky (A)	5	75	16
Saving / Borrowing (S, B)	25	375	16
Full information (FI)	25	11,625	17
Moral hazard (MH)	25	11,625	23
Limited commitment (LC)	25	11,625	32
Hidden output (HO)	25	11,625	77
Unobserved investment (UI), stage 1	250	1,650	122
Unobserved investment (UI), stage 2	550	8,370	2,507
Unobserved investment (UI), total	137,500	n.a.	n.a.

Note: This table assumes the following grid sizes that used in the estimation: #Q=5, #K=5, #Z=3, #B=5, #T=31; #W=5; and #**W**=50 and #**W**m=110 for the UI model

Table 2 - Variable Grids Used in the Estimation

Variable	grid size (number of points)	grid range
income/cash flow, Q	5	[.04,1.75] from data percentiles
business assets, K	5	[0, 1] from data percentiles
effort, Z	3	[.01, 1]
savings/debt, B	5 (6 for B regime)	S: [-2, 0], B: [-2, .82]
transfers/consumption, C	31 for MH/FI/LC, endog. for B/S/A	[.001, 0.9]
promised utility, W	5	endogenous

Table 3 - Thai data summary statistics

	Rural data, 1999-2005	Urban data, 2005-2009
Consumption expenditure, c		
mean	64.172	148.330
standard deviation	53.284	131.710
median	47.868	115.171
Income, q		
mean	128.705	635.166
standard deviation	240.630	1170.400
median	65.016	361.000
Business assets, k		
mean	80.298	228.583
standard deviation	312.008	505.352
median	13.688	57.000
Investment, i		
mean	6.249	17.980
standard deviation	57.622	496.034
median	0.020	1.713

1. Sample size in the rural data is 531 households observed over seven consecutive years (1999-2005).
2. Sample size in the urban data is 475 households observed over five consecutive years (2005-2009).
3. All summary statistics in the Table are computed from the pooled data. Units are '000s Thai baht.

Table 4 - Parameter Estimates using 1999-00 Thai Rural Data

Business assets, investment and income, (k,i,q) data

Model	γ_{me}	σ	θ	$\mu_{w/b}$[1]	$\gamma_{w/b}$	LL Value[2]
Moral hazard - MH	0.1632	0.0465	1.3202	0.4761	0.0574	**-3.1081**
	(0.0125)	*(0.0000)*	*(0.0000)*	*(0.0139)*	*(0.0005)*	
Full information - FI	0.1625	0.0323	1.1928	0.4749	0.0591	**-3.1100**
	(0.0132)	*(0.0060)*	*(0.0770)*	*(0.0351)*	*(0.0138)*	
Limited commitment - LC	0.1487	3.8032	0.6210	0.9723	0.0713	**-3.1166**
	(0.0081)	*(0.2337)*	*(0.1756)*	*(0.0083)*	*(0.0001)*	
Borrowing & Lending - B	0.0950	4.2990	0.1091	0.8883	0.0065	**-2.5992**
	(0.0059)	*(0.0880)*	*(0.0000)*	*(0.0269)*	*(0.0153)*	
Saving only - S *	0.0894	5.7202	9.2400	0.9569	0.0101	**-2.5266**
	(0.0068)	*(0.0000)*	*(0.0000)*	*(0.0087)*	*(0.0075)*	
Autarky - A	0.1203	3.1809	9.2000	n.a.	n.a.	**-2.7475**
	(0.0046)	*(0.6454)*	*(0.0000)*	n.a.	n.a.	

Consumption and income, (c,q) data

Model	γ_{me}	σ	θ	$\mu_{w/b}$	$\gamma_{w/b}$	LL Value
Moral hazard - MH *	0.1324	0.5020	1.9248	0.5499	0.0514	**-0.9472**
	(0.0114)	*(0.0000)*	*(0.0000)*	*(0.0053)*	*(0.0005)*	
Full information - FI	0.1528	0.6450	8.8301	0.6805	0.1169	**-1.0223**
	(0.0087)	*(0.0000)*	*(0.0000)*	*(0.0048)*	*(0.0025)*	
Limited commitment - LC	0.1291	2.7560	0.3732	0.0005	0.4290	**-1.0549**
	(0.0120)	*(0.0895)*	*(0.0973)*	*(0.0358)*	*(0.0310)*	
Borrowing & Lending - B	0.1346	4.3322	1.8706	0.8397	0.0311	**-1.0558**
	(0.0130)	*(0.0197)*	*(0.0000)*	*(0.0045)*	*(0.0004)*	
Saving only - S *	0.1354	2.9590	0.0947	0.9944	0.0516	**-1.0033**
	(0.0074)	*(0.0343)*	*(0.8556)*	*(0.0133)*	*(0.0180)*	
Autarky - A	0.1769	1.2000	1.2000	n.a.	n.a.	**-1.1797**
	(0.0087)	*(0.0000)*	*(4.2164)*	n.a.	n.a.	

Business assets, consumption, investment, and income, (c,q,i,k) data

Model	γ_{me}	σ	θ	$\mu_{w/b}$	$\gamma_{w/b}$	LL Value
Moral hazard - MH	0.1581	0.0342	0.9366	0.3599	0.0156	**-2.8182**
	(0.0073)	*(0.0000)*	*(0.0000)*	*(0.0013)*	*(0.0010)*	
Full information - FI	0.1434	0.1435	1.0509	0.5608	0.1244	**-2.8119**
	(0.0083)	*(0.0018)*	*(0.0009)*	*(0.0112)*	*(0.0105)*	
Limited commitment - LC	0.3061	3.0695	8.0000	0.3834	0.0477	**-4.0867**
	(0.0057)	*(0.0230)*	*(1.5353)*	*(0.0272)*	*(0.0176)*	
Borrowing & Lending - B	0.1397	1.0831	8.1879	0.9571	0.0398	**-2.5582**
	(0.0071)	*(0.1102)*	*(0.2536)*	*(0.0359)*	*(0.0267)*	
Saving only - S *	0.1245	5.6697	0.1114	0.9839	0.0823	**-2.3825**
	(0.0077)	*(0.0225)*	*(0.0744)*	*(0.0248)*	*(0.0432)*	
Autarky - A	0.1394	1.6922	9.2000	n.a.	n.a.	**-2.6296**
	(0.0050)	*(0.3157)*	*(0.0000)*	n.a.	n.a.	

1. $\mu_{w/b}$ and $\gamma_{w/b}$ (the mean and standard deviation of the w or b initial distribution) are reported relative to the variables' grid range

2. Normalized (divided by n) log-likelihood values;

3. Bootstrap standard errors are in parentheses below each parameter estimate.

* denotes the best fitting regime (including ties)

Table 5 - Model Comparisons[1,2] using Thai Rural Data - Baseline Vuong Test Results

Comparison	MH v FI	MH v LC	MH v B	MH v S	MH v A	FI v LC	FI v B	FI v S	FI v A	LC v B	LC v S	LC v A	B v S	B v A	S v A	Best Fit
1. Using (k,i,q) data																
1.1 years: 1999-00	MH*	tie	B***	S***	A***	tie	B***	S***	A***	B***	S***	A***	S***	B***	S***	S
1.2 years: 2004-05	FI***	MH***	B***	S***	A***	FI**	B***	S***	A***	B***	S***	A***	tie	B***	S***	B,S
2. Using (c,q) data																
2.1 year: 1999	MH***	MH**	MH**	tie	MH***	FI*	tie	tie	FI***	tie	tie	LC**	S***	B***	S***	MH,S
2.2 year: 2005	tie	MH***	tie	tie	tie	FI***	tie	S***	tie	B**	S***	tie	S**	tie	S***	S,MH
3. Using (c,q,i,k) data																
3.1 years: 1999-00	tie	MH***	B***	S***	A**	FI***	B***	S***	A**	B***	S***	A***	S***	tie	S***	S
3.2 years: 2004-05	FI***	MH***	B***	S***	A***	FI***	B***	S***	A**	B***	S***	A***	S***	tie	S**	S
4. Two-Year Panel																
4.1 (c,q) data, years: 1099 and 00	MH***	MH***	B***	S***	MH**	FI**	B***	S***	tie	B***	S***	tie	tie	B***	S***	S,B
4.2 (c,q) data, years: 1999 and 05	MH***	MH***	tie	tie	MH***	FI***	B***	S***	tie	B***	S***	tie	tie	B***	S***	B,S,MH
5. Dynamics																
5.1 99 k distribution & 04-05 (c,q,i,k)	FI***	MH***	B***	tie	tie	FI***	B***	tie	FI*	B***	S***	A***	B***	B***	S**	B
5.2 99 k distribution & 05 (c,q)	tie	MH***	tie	tie	MH***	FI***	tie	tie	FI***	B***	S***	A***	tie	B***	S***	S,B,FI,MH
5.3 99 k distribution & 04-05 (k,i,q)	FI***	LC***	B***	S**	MH**	tie	B***	S*	FI**	B***	S*	LC**	B***	B***	S***	B

Table 6 - Model Comparisons[1] using Thai Rural Data - Networks

Comparison	MH v FI	MH v LC	MH v B	MH v S	MH v A	FI v LC	FI v B	FI v S	FI v A	LC v B	LC v S	LC v A	B v S	B v A	S v A	Best Fit
1. Networks by friend/relative																
1.1 (c,q) data, in network, n=391	MH***	MH***	MH***	MH*	MH***	FI*	tie	tie	FI**	tie	tie	LC**	S***	B***	S***	MH
1.2 (k,i,q) data, in network	tie	tie	B***	S***	A***	FI**	B***	S***	A***	B***	S***	A***	S**	B**	S***	S
1.3 (c,q,i,k) data,in network	tie	MH***	B***	S***	A**	FI***	B***	S***	A***	B***	S***	A***	S***	tie	S**	S
1.4 (c,q) data, not in network	tie	MH***	tie	tie	tie	FI*	tie	tie	tie	tie	tie	tie	tie	B*	tie	all tied
1.5 (c,q,i,k) data, not in network	tie	MH***	tie	S***	tie	FI***	tie	S***	A**	B***	S***	A***	S***	tie	S*	S
2. Networks by gift or loan																
2.1 (c,q) data, in network, n=357	FI**	MH***	MH**	tie	MH***	FI***	FI***	FI**	FI***	tie	S***	LC*	S***	B***	S***	FI
2.2 (k,i,q) data, in network	tie	tie	B***	S***	A***	tie	B***	S***	A***	B***	S***	A***	S**	B**	S***	S
2.3 (c,q,i,k) data, in network	tie	MH***	B***	S***	A**	FI***	B***	S***	A**	B***	S***	A***	S***	tie	S**	S
2.4 (c,q) data, not in network	tie	MH***	MH**	tie	MH**	FI*	FI***	tie	FI*	tie	tie	tie	S***	tie	S*	MH,FI,S
2.5 (c,q,i,k) data, not in network	tie	MH***	B***	S***	tie	FI***	B***	S***	tie	B***	S***	A***	S***	tie	S***	S

Notes: 1. *** = 1%, ** = 5%, * = 10% two-sided significance level, the better fitting model abbreviation is displayed; 2. Vuong statistic cutoffs: >2.575 = ***; >1.96 = **; >1.645 = *; <1.645 = "tie"

Table 7 - Model Comparisons[1,2] using Thai Urban Data - Vuong Test Results

Comparison	MH v FI	MH v LC	MH v B	MH v S	MH v A	FI v LC	FI v B	FI v S	FI v A	LC v B	LC v S	LC v A	B v S	B v A	S v A	Best Fit
1. Using (c,q,i,k) data																
1.1. years: 2005-06	MH***	MH***	MH***	MH***	MH***	FI***	B***	S***	FI*	B***	S***	A***	S***	B***	S***	MH
1.2. years: 2008-09	MH***	MH***	MH***	MH***	MH***	FI***	B***	S***	tie	B***	S***	A***	S***	B***	S***	MH
2. Using (c,q) data																
2.1. year: 2005	tie	MH**	MH***	MH**	MH***	tie	FI***	FI**	FI***	LC***	tie	LC***	S***	B***	S***	MH,FI
2.2. year: 2009	MH*	MH***	tie	MH*	MH***	FI***	tie	tie	FI***	B***	S***	A***	tie	B***	S***	MH,B
3. Using (k,i,q) data																
3.1. years: 2005-06	tie	MH***	tie	S***	tie	FI**	tie	S***	tie	B***	S***	tie	S***	tie	S**	S
3.2. years: 2008-09	FI*	tie	B***	S***	A***	FI**	B***	S***	tie	B***	S***	A**	tie	tie	S*	S,B
4. Two-year panel																
4.1. (c,q) data, years: 2005 and 06	tie	MH***	MH***	tie	MH***	FI***	FI***	tie	FI***	tie	S***	tie	S***	B**	S***	S,MH,FI
4.2. (c,q) data, years: 2005 and 09	MH***	MH***	MH***	MH***	MH***	FI***	FI***	FI***	FI***	LC***	tie	LC***	S***	B***	S***	MH

Table 8 - Model Comparisons[1,2] using Thai Rural Data and Estimated production function

Comparison	MH v FI	MH v LC	MH v B	MH v S	MH v A	FI v LC	FI v B	FI v S	FI v A	LC v B	LC v S	LC v A	B v S	B v A	S v A	Best Fit
1. Using (k,i,q) data																
1.1 years: 99-00	FI**	LC***	B***	S***	A***	LC***	B***	S***	A***	B***	S***	A***	S*	B***	S***	S
1.2 years: 04-05	MH***	tie	B***	S***	A***	LC***	B***	S***	A***	B***	S***	A***	S***	tie	S**	S
2. Using (c,q) data																
2.1 year: 99	MH*	MH***	tie	tie	MH***	FI***	B***	S***	FI*	B***	S***	tie	tie	B***	S***	B,S,MH
2.2 year: 05	MH**	MH**	B***	S***	tie	tie	B***	S***	A**	B***	S***	A***	B**	B**	tie	B
3. Using (c,q,i,k) data																
3.1 years: 99-00	tie	MH***	B***	S***	A***	FI***	B***	S***	A***	B***	S***	A***	tie	B***	S***	B,S
3.2 years: 04-05	MH***	LC***	B***	S***	A***	LC***	B***	S***	A***	B***	S***	A***	S*	B**	S***	S
4. Two-year panel																
4.1. (c,q), years: 99 and 00	MH***	MH**	tie	S**	MH***	LC**	B***	S***	FI***	B**	S***	LC***	tie	B***	S***	S,B
4.2. (c,q), years: 99 and 05	MH*	MH***	tie	tie	MH***	tie	tie	tie	FI***	B*	tie	LC***	tie	B***	S***	B,MH,S,FI

Notes: 1. *** = 1%, ** = 5%, * = 10% two-sided significance level, the better fitting model abbreviation is displayed; 2. Vuong statistic cutoffs: >2.575 = ***; >1.96 = **; >1.645 = *; <1.645 = "tie"

Table 9 - Model Comparisons[1] using Thai Rural Data - Robustness Runs

Comparison	MH v FI	MH v LC	MH v B	MH v S	MH v A	FI v LC	FI v B	FI v S	FI v A	LC v B	LC v S	LC v A	B v S	B v A	S v A	Best Fit
1. Risk neutrality[2]																
1.1 (c,q) data	MH***	MH***	MH***	MH***	MH***	LC***	B***	S***	A***	B***	S***	A***	S**	tie	S***	MH
1.2 (k,i,q) data	tie	tie	B***	S***	A***	FI**	B***	S***	A***	B***	S***	A***	B***	B***	tie	B
1.3 (c,q,i,k) data	tie	tie	B***	S***	A***	LC**	B***	S***	A***	B***	S***	A***	tie	B*	S***	S,B
2. Fixed measurement error variance																
2.1 (c,q) data	tie	MH***	MH***	tie	MH***	FI***	FI***	tie	FI***	tie	S***	tie	S***	B***	S***	MH,S,FI
2.2 (k,i,q) data	tie	MH***	B***	S***	A***	FI***	B***	S***	A***	B***	S***	A***	S***	B***	S***	S
2.3 (c,q,i,k) data	FI***	MH***	B***	S***	A***	FI***	B***	S***	A*	B***	S***	A***	S***	tie	S***	S
3. Investment adjustment costs																
3.1. (c,q) data	MH**	MH***	B**	tie	MH***	FI***	B***	S**	tie	B***	S***	tie	B*	B***	S***	B
3.2 (k,i,q) data	tie	tie	B**	S***	A***	tie	B***	S***	A***	B***	S***	A***	S*	A*	tie	S,A
3.3 (c,q,i,k) data	tie	MH***	tie	S**	MH**	FI***	tie	tie	FI***	B***	S***	A***	S**	B***	S***	S,FI
4. Removed fixed effects																
4.1 removed year fixed effects, cqik	tie	MH***	B***	S***	A***	FI***	B***	S***	A***	B***	S***	A***	S*	tie	S*	S
4.2 removed fixed effects (yr+hh), kiq	tie	tie	B*	S***	A***	tie	B*	S***	A***	B*	S***	A***	S***	A***	S*	S
4.3 removed fixed effects (yr+hh), cq	MH*	MH***	MH***	MH***	MH***	FI***	FI***	FI**	FI***	LC***	S**	LC***	S***	B***	S***	MH
4.4 removed fixed effects (yr+hh), cqik	MH***	MH***	MH***	MH***	MH***	FI***	FI***	FI***	FI***	LC***	S***	LC***	S***	B***	S***	MH
4.5 removed fixed effects, estim. pr. f-n	FI***	tie	tie	tie	MH***	FI***	tie	tie	FI***	tie	S*	LC***	tie	B***	S***	S,B,FI,MH
5. Other robustness runs (with 1999-00 c,q,i,k data unless otherwise indicated)																
5.1 alternative assets definition	tie	tie	MH**	S***	tie	tie	FI**	S***	tie	tie	S***	tie	S***	A***	tie	S
5.2 alternative interest rate, R=1.1	tie	MH***	B***	S***	A*	FI***	B***	S***	A*	B***	S***	A***	tie	B***	S***	S,B
5.3 alternative depreciation rate, δ=0.1	FI***	MH***	B***	S***	A***	FI***	B***	S***	A**	B***	S***	A***	tie	B*	S***	S,B
5.4 coarser grids	MH***	MH***	B***	S***	A***	FI***	B***	S***	A***	B***	S***	A***	B**	B***	S***	B
5.5 denser grids	MH***	MH***	B***	S***	A***	FI***	B***	S***	A***	B***	S***	A***	tie	B***	S***	B,S
6. Runs with hidden output (HO) and unobserved investment (UI) models[3]																
	v MH	**v FI**	**v B**	**v S**	**v A**	**v LC**										
6.1 hidden output model, (c,q,i,k)	tie	tie	B***	S***	A***	HO***										B,S
6.2 unobserved investment model, (c,q,i,k)	UI***	UI***	B***	S***	tie	UI***										B

1. *** = 1%, ** = 5%, * = 10% Vuong (1989) test two-sided significance level. Listed is the better fitting model or "tie" if the models are tied. Sample size is n=531; data are for 1999-00 unless noted otherwise.

2. The upper bound of the output grid, Q was adjusted to 1.25 for these runs, since our baseline grid produced no solution for the LC regime for σ = 0.

3. For computational reasons the HO model is computed with estimated production function (read with table 6a); the UI model is with coarser grids (read with line 6.4).

Table 10 - Parameter Estimates using Simulated Data from the Moral Hazard (MH) Model

Assets, investment and income, (k,i,q) data

Model	γ_{me}	σ	θ	ρ	$\mu_{w/b}$[1]	$\gamma_{w/b}$	LL Value[2]
Moral hazard - MH *	0.0935	0.6557	0.1000	0.2212	0.8289	0.0778	**-1.0695**
Full information - FI *	0.0937	0.5495	0.1000	0.2720	0.8111	0.1078	**-1.0692**
Limited commitment - LC	0.1053	1.3509	1.1087	-4.2141	0.4483	0.5468	**-1.2410**
Borrowing & Lending - B	0.1011	1.0940	1.0811	-1.5783	0.0096	0.9995	**-1.1821**
Saving only - S	0.0972	0.5000	1.2043	-1.8369	0.5184	0.1697	**-1.1407**
Autarky - A	0.2927	0.0000	2.0000	2.2117	n.a.	n.a.	**-2.5390**
baseline parameters	*0.1000*	*0.5000*	*2.0000*	*0.0000*	*0.5000*	*0.3500*	

Consumption and income, (c,q) data

Model	γ_{me}	σ	θ	ρ	$\mu_{w/b}$	$\gamma_{w/b}$	LL Value
Moral hazard - MH *	0.1041	0.4851	2.7887	-0.2338	0.4780	0.2867	**-0.1462**
Full information - FI	0.1102	0.4462	0.0934	-1.2892	0.5056	0.2644	**-0.1784**
Limited commitment - LC	0.1157	1.1782	1.2024	-10.9857	0.2276	0.6321	**-0.2185**
Borrowing & Lending - B	0.1160	0.6007	0.1544	-1.5090	0.5202	0.3489	**-0.2182**
Saving only - S	0.1077	0.0000	1.9849	3.0075	0.4204	0.4527	**-0.1842**
Autarky - A	0.1868	0.0276	0.9828	0.2036	n.a.	n.a.	**-0.7443**
baseline parameters	*0.1000*	*0.5000*	*2.0000*	*0.0000*	*0.5000*	*0.3500*	

Assets, consumption, investment, and income, (c,q,i,k) data

Model	γ_{me}	σ	θ	ρ	$\mu_{w/b}$	$\gamma_{w/b}$	LL Value
Moral hazard - MH *	0.0952	0.5426	2.1951	0.2267	0.5005	0.3464	**-0.8952**
Full information - FI	0.1358	0.5436	0.0967	-6.4718	0.5567	0.2862	**-1.4184**
Limited commitment - LC	0.1381	1.2000	0.1239	-36.3392	0.2654	0.5952	**-1.4201**
Borrowing & Lending - B	0.1339	1.2000	7.7164	-3.0189	0.4048	0.3238	**-1.5624**
Saving only - S	0.1678	0.0000	0.0727	-1.1738	0.3818	0.2771	**-1.7803**
Autarky - A	0.3302	1.2000	0.1000	0.4681	n.a.	n.a.	**-3.0631**
baseline parameters	*0.1000*	*0.5000*	*2.0000*	*0.0000*	*0.5000*	*0.3500*	

1. $\mu_{w/b}$ and $\gamma_{w/b}$ (the mean and standard deviation of the w or b initial distribution) are reported relative to the variables' grid range

2. Normalized (divided by n) log-likelihood values;

* denotes the best fitting regime (including tied)

All runs use data with sample size n=1000 generated from the MH model at the *baseline parameters*

Table 11 - Model Comparisons using Simulated Data[1] - Vuong Test Results

Comparison	MH v FI	MH v LC	MH v B	MH v S	MH v A	FI v LC	FI v B	FI v S	FI v A	LC v B	LC v S	LC v A	B v S	B v A	S v A	Best Fit
1. Using (k,i,q) data																
1.1 low measurement error	tie	MH***	MH***	MH***	MH***	FI***	FI***	FI***	FI***	LC***	LC***	LC***	S**	B***	S***	MH,FI
1.2 high measurement error	tie	tie	tie	tie	MH***	tie	B**	tie	FI***	tie	tie	LC***	tie	B***	S***	all but A
2. Using (c,q) data																
2.1 low measurement error	MH***	MH***	MH***	MH***	MH***	FI***	FI**	tie	FI***	tie	S*	LC***	S**	B***	S***	MH
2.2 high measurement error	FI***	tie	B*	MH*	MH***	tie	tie	FI***	FI***	tie	tie	LC***	B***	B***	S***	B,FI
3. Using (c,q,i,k) data																
3.1 low measurement error	MH***	MH***	MH***	MH***	MH***	tie	FI***	FI***	FI***	LC***	LC***	LC***	B***	B***	S***	MH
3.2 high measurement error	tie	MH***	MH***	MH***	MH***	FI***	FI***	FI***	FI***	LC**	LC***	LC***	B***	B***	S***	MH,FI
4. Two-year (c,q) panel, t = 0, 1																
4.1 low measurement error	MH***	MH***	MH***	MH***	MH***	FI***	FI***	FI***	FI***	LC***	LC***	LC***	B***	B***	S***	MH
4.2 high measurement error	tie	tie	MH***	MH***	MH***	tie	FI***	FI***	FI***	LC***	LC***	LC***	B***	B***	S***	MH,FI,LC
5. Robustness runs with simulated data[2]																
5.1 (c,q) data long panel (t = 0, 50)	MH***	MH***	MH***	MH***	MH***	FI***	FI***	FI***	FI***	LC***	LC***	LC***	B***	B***	S***	MH
5.2 zero measurement error	MH***	MH***	MH***	MH***	MH***	FI***	tie	FI*	FI***	B*	tie	LC***	B***	B***	S***	MH
5.3 sample size n = 200	MH***	MH***	MH***	MH***	MH***	tie	tie	FI***	FI***	tie	LC***	LC***	B***	B***	S***	MH
5.4 sample size n = 5000	MH***	MH***	MH***	MH***	MH***	tie	FI***	FI***	FI***	LC***	LC***	LC***	B***	B***	S***	MH
5.5 coarser grids	MH***	MH***	MH***	MH***	MH***	FI***	FI***	FI***	FI***	LC***	LC***	LC***	B***	B***	S***	MH
5.6 denser grids	MH***	MH***	MH***	MH***	MH***	FI**	FI***	FI***	FI***	LC***	LC***	LC***	B**	B***	S***	MH
5.7 heterogeneous productivity	MH***	MH***	MH***	MH***	MH***	tie	tie	FI***	FI***	tie	LC***	LC***	B***	B***	S***	MH
5.8 heterogeneous risk-aversion	MH***	MH***	MH***	MH***	MH***	FI**	FI***	FI***	FI***	LC***	LC***	LC***	B***	B***	S***	MH

1. *** = 1%, ** = 5%, * = 10% two-sided significance level, the better fitting model regime's abbreviation is displayed. Data-generating model is MH and sample size is n = 1000 unless stated otherwise.

2. these runs use (c,q,i,k) data simulated from the MH model and low measurement error ($\gamma_{me} = 0.1$) unless stated otherwise

Table 13: Consumption Euler equation GMM test as in Ligon (1998), rural sample

Instruments	b	std. error	[95% conf. interval]		J-test
---	-0.3358*	0.0602	-0.454	-0.218	n.a.
income	-0.3257*	0.0546	-0.433	-0.219	1.006
income, capital	-0.3365*	0.0499	-0.434	-0.239	2.389
income, capital, avg. consumption	-0.3269*	0.0492	-0.423	-0.231	2.793

Notes:
1. b is the estimate of the risk aversion coefficient; assuming households are risk-averse,
a negative b suggests the correct model is B (standard EE); a positive b suggests MH (inverse EE)
2. the estimates are obtained using continuous updating GMM (Hansen, Heaton and Yaron, 1996).
Matlab code adapted from K. Kyriakoulis, using HACC_B method with optimal bandwidth.

Table 14: Investment Euler equation GMM test as in Bond and Meghir (1994), rural sample

Dynamic panel-data estimation, one-step difference GMM using lags of 2 or more for instruments

Group variable: household	Number of observations: 1552
Time variable : year	Number of groups: 388
Number of instruments = 24	Observations per group: 4

dependent variable = i_t / k_t	Coef	Robust st. err.	z	P > |z|	[95% conf. interval]	
i_{t-1} / k_{t-1}	0.3232775	0.0595142	5.43	0.000	0.2066317	0.4399232
$(i_{t-1} / k_{t-1})^2$	-0.096548	0.2777705	-0.35	0.728	-0.6409683	0.4478719
q_{t-1} / k_{t-1}	0.0002172	0.0002812	0.77	0.440	-0.0003339	0.0007683
year dummies	included					

Arellano-Bond test for AR(1) in first differences: z = -1.87 Pr > z = 0.061
Arellano-Bond test for AR(2) in first differences: z = -0.48 Pr > z = 0.628
Arellano-Bond test for AR(3) in first differences: z = 1.25 Pr > z = 0.211
Hansen test of overid. restrictions: chi2(17) = 22.29 Prob > chi2 = 0.174

Note: observations with zero assets (k) were excluded.