August 23, 2005

Financial Deepening, Inequality, and Growth: A Model-Based Quantitative Evaluation

Robert M. Townsend and Kenichi Ueda *

ABSTRACT

We propose a coherent unified approach to the study of the linkages among economic growth, financial structure, and inequality, bringing together disparate theoretical and empirical literature. That is, we show how to conduct model-based quantitative research on transitional paths. With analytical and numerical methods, we calibrate and make tractable a prototype canonical model and take it to an application, namely, Thailand 1976–96, an emerging market economy in a phase of economic expansion with uneven financial deepening and increasing inequality. We look at the expected path generated by the model and conduct robustness experiments. Because the actual path of the Thai economy is imagined here to be just one realization of many possible histories of the model economy, we construct a covariance-normalized squared error metric of closeness and find the best-fit simulation. We also construct a confidence region from a set of simulations and formally test the model. We broadly replicate the actual data and identify anomalies.

---

*Robert M. Townsend is a professor at the University of Chicago and a consultant at the Federal Reserve Bank of Chicago. Kenichi Ueda is an economist at the International Monetary Fund. The views expressed in this paper are those of authors and do not necessarily represent those of the FRB Chicago, the Federal Reserve System, or the IMF. We thank the National Statistics Office of Thailand for permission to use the SES data. We received helpful comments from Robin Brooks, Fabio Canova, Marco Cagetti, Larry Christiano, Angus Deaton, Paul Evans, Jess Gaspar, Hyeok Jeong, Ashok Mody, Tomoyuki Nakajima, Hajime Takizawa, Chris Udry, Annette Vissing-Jørgensen and the participants of the NBER general equilibrium conference at New York University, the conference honoring Rolf R. Mantel in Argentina, seminars at Cornell University, GRIPS, Hitotsubashi University, Ohio State University, Oxford University, and the IMF, as well as the IMF conference on Macroeconomic Policy and Poverty Reduction. We are especially indebted to the editors and referees. Townsend acknowledges financial support from the NSF and NICHD. We thank Natalia Ramondo for excellent research assistance.

JEL Classification Numbers: C52, G20, O16, O41
Keywords: Financial deepening, inequality, growth, transitional dynamics, stochastic calibration, confidence region.

I. Introduction

We propose a coherent unified approach to the study of the linkages among economic growth, financial structure, and inequality. Of course, the relationship between financial structure and economic growth has long been studied both empirically and theoretically. Yet, on the one hand, empirical studies have been mainly focused on statistical relationships without a serious study of underlying mechanisms that generate the observations. On the other hand, most theoretical studies have depicted clean but simple mechanisms without serious consideration given to the models' quantitative predictions. The same dichotomy between theory and empirical work exists in the literature on inequality and growth.

Early seminal empirical contributions focusing on growth and financial structure are Goldsmith (1969), Shaw (1973), and McKinnon (1973). A more recent empirical treatment is King and Levine (1993). This body of empirical work establishes that financial deepening is at least an intrinsic part of the growth process and may be causal—that is, repressed financial systems harm economic growth. Theoretical efforts at modeling growth and endogenous financial deepening include Townsend (1978, 1983) and Greenwood and Jovanovic (1990) (hereinafter referred to as GJ). These models posit costly bilateral exchange or intermediation costs—for example, a fixed cost to enter the formal financial system and marginal costs to subsequent transactions. Other theoretical contributions such as Bencivenga and Smith (1991) turn intermediation on and off exogenously and have an external effect that makes growth with intermediation higher. Saint-Paul (1992) features limited diversification and multiple equilibrium growth paths, some with developed financial systems and specialized technologies and others with the opposite. In turn, Acemoglu and Zilibotti (1997) show that capital accumulation is associated with increasing intermediation and that better diversification, which comes with higher levels of wealth, reduces the variability of growth.

Likewise well known are seminal contributions on growth and inequality. Kuznets (1955) posited that growth is associated with increasing and eventual decreasing inequality. Interest and controversies, especially with respect to cross-country regressions, have continued ever since. A recent paper, Forbes (2000), confirms previous regression studies that high (initial) inequality is associated with low subsequent long-run growth but finds that the relationship is the opposite for the medium term. Resting separately from this strand of the empirical literature are the deservedly well-known theoretical contributions more motivated by Kuznets's original assertion that growth may bring increasing, and eventually decreasing, inequality—namely, Aghion and Bolton (1997), Piketty (1997), Banerjee and Newman (1993), and Lloyd-Ellis and Bernhardt (2000).

We have a concern about this dichotomy between theories and empirical studies. Although most of the theoretical models characterize economic growth with financial deepening and changing inequality as transitional phenomena, typical empirical research employs regression analysis to find a coefficient capturing the effect of financial depth or inequality on growth. The implicit assumptions of stationarity and linearity are incorrect, even after taking logs and lags, if the variables of actual economies lie on complex transitional growth paths, as they do in the theoretical models. Using artificial data generated by a canonical model that *by construction* displays transitional growth with financial deepening and increasing inequality, we can sometimes replicate the typical empirical results of the literature: financial deepening appears to lead to subsequent higher growth, and inequality to subsequent lower growth. But the statistical significance is weak and sensitive to initial history, time frame, the inclusion of covariates, and so on. Evidently regression coefficients are not informative about the underlying true relationships. In pointing this out, we add to the list of concerns which have been raised in recent literature—for example, in Banerjee and Duflo (2000).

Taking a more constructive tack, we show how to conduct quantitative research on transitional growth paths—that is, how to test a model and learn something about actual economies from potential rejections. Our canonical model is based on GJ.[2] As a prerequisite for a numerical study, we need to characterize analytical properties as much as possible. GJ characterizes primarily the initial and asymptotic economy but leaves the all-important transitions somewhat unclear. GJ also studies only the log utility function and does not take its characterization of the log case to data. Here we extend the model to include a wider class of CRRA utility functions. We characterize much of the transitional dynamics analytically and, in doing so, provide new results. The seemingly nonconvex technology of participation is shown, under some conditions, to be convexified by the optimal choice of portfolio shares between risky and safe assets. Consequently, savings and portfolio choice are uniquely determined depending on the wealth level. In particular, ironically, those risk-averse households and businesses without access are not condemned to low yield (though safe) technologies but rather shift toward risky enterprizes, especially as their wealth approaches a critical value. These make clear the rich and potentially complicated dynamics not obvious in the original GJ formulation. Indeed, the single-valuedness of savings and portfolio choice facilitates further research into transitional dynamics using numerical methods and we find, for example, overall inequality movement is not necessarily monotonic on the growth path—it can increase, then decrease, and then increase again as it moves slowly toward its asymptotic steady state. Financial deepening and growth are not monotonic either.

With the model made tractable, we take it to an application, namely, Thailand between 1976–96, an emerging market economy that was in a phase of economic expansion with uneven financial deepening and increasing (and then decreasing) inequality. The Thai economy serves as a prototypical example of the growth and inequality phenomena, pervasive in other countries, that motivate the growth, financial deepening, and inequality literature. Using the Thai *Socio-Economic Survey* (SES), Jeong (2000) finds that growth and inequality are strongly associated with financial deepening. We emphasize, however, that our methods are not peculiar to Thailand and we hope to extend the analysis to other countries.

In the spirit of the business cycle literature, we calibrate the parameter values. Thus, the benchmark parameters are set from several sources. Data on the yields of relatively safe assets or occupations—5.4 percent per year for agriculture—and idiosyncratic shocks for business come from

[2]There are a few other model-based contributions to empirical work on growth and wealth inequality. Álvarez and Díaz (2002) study evolution of wealth inequality in a nonstochastic neo-classical growth model with minimum consumption requirements and apply it to the U.S. economy. It is a calibration study of growth and inequality of wealth, but the growth rate is not affected by inequality of wealth because of perfect capital markets and identical incomes among households. Likewise De Nardi (2004) focuses on savings and bequests to explain wealth inequality in the U.S. and Sweden. De Nardi's is a steady-state calibration exercise based on an overlapping generation model. This also contains an excellent review of the inter-generational inequality literature.

the Townsend-Thai data.[3] Risk aversion is set at values typically found in the financial economics literature, and the real rate of interest implied by a preference for current consumption is set at 4 percent. Aggregate shocks are difficult to pin down, and we take advantage of the average and the variation of the observed GDP growth rate over the sample period. The marginal cost of utilizing the financial system is set at low values, but the higher fixed cost of entry is such that, as in the Thai *Socio-Economic Survey* (SES), 6 percent of the Thai population would have had access to the financial system in 1976, using a distribution of wealth estimated from the same 1976 SES data. The model is simulated at these and nearby values to deliver predicted paths, which can be compared to the actual Thai data.

Note that we are calibrating a non-steady-state stochastic model. The business cycle literature calibrates around stochastic steady states, not transitions. For example, Krusell and Smith (1998) find that a pseudo representative-agent model can almost perfectly characterize the behavior of the macroeconomic aggregates of an actual incomplete market economy with wealth heterogeneity among agents. This is not the case here. This different outcome stems from the fact that they look at business cycle implications in a stochastic steady state and also their model does not contain a source of non-linearity, such as the fixed cost to participate the financial system in our model. There is also a literature on transitions or out-of-steady-state dynamics devoted to the study of depressions, but the models are deterministic (e.g., Hayashi and Prescott, 2002). Indeed, it is a challenging task to calibrate transitional paths in a stochastic environment, because there is no precedent sensible statistic criteria for goodness of fit. Here, we examine the fit of the model in three ways: the expected path, the best-fit prediction, and the confidence region.

First, we look at the expected path of the model. It is broadly consistent with the actual pattern of growth with increasing inequality along with financial deepening. However, although the computed expected participation rate and the Theil index (an inequality measure) in the model almost trace out a smoothed version of the actual Thai data, the computed expected GDP growth rate is lower than the actual Thai path. We vary the key parameters and conduct robustness experiments, exploring some tradeoffs.

Second, we look at the best-fit path among 1000 simulations. Because the actual path of the Thai economy is imagined here to be just one realization of many possible histories of the model economy, the actual Thai path should differ from the expected path of the model. We construct a metric of closeness between a simulated path and the actual data, considering the growth rate, financial deepening, and inequality variables over the entire 1976–96 period to be one particular realization. We pick the best-fit simulation under this metric. The best-fit simulation shows a reasonable match with GDP growth rate, but misses a sharp upturn of the financial deepening in the mid-1980s as well as the eventual downturn of inequality in 1990s.

Finally, we examine whether the actual Thai data lie within a confidence region generated by the model, using a covariance-normalized mean squared error criteria constructed solely from the model-generated histories. The model imposes sharp restrictions on the data, and indeed, at the benchmark parameter values, the model is rejected. But this guides us in identifying where the model fits well and where it does not.

[3]See Townsend and others (1997).

Of course, we are not unaware of the tension in this paper between the working out of the details of a structural and well articulated but highly abstract version of reality, on the one hand, and its serious application to data on the other. But the goal here is to show that these two pieces can be brought together. This, we believe, is the way to make progress toward understanding the true relationships among endogenous financial deepening, economic development, and changing inequality.

II. MODEL

A. Notation

The model is a simple, tractable growth model with a financial sector. It is assumed that a large number of people live in the economy and they are consumers and entrepreneurs at the same time. More specifically, there is a continuum of agents in the economy as if with names indexed on the interval $[0, 1]$. At the beginning of each period, they start with their assets k_t. After they consume c_t, some portion of the assets, they use the remaining assets (i.e., savings s_t) to engage in productive activities.

An individual can engage in two types of productive activities: a safe but low-return occupation and high-risk high-return business. The safe projects are assumed to return δ and the risky businesses are assumed to return $\theta_t + \epsilon_t$, where $\theta_t \in \Theta$ is a shock common to all the people and $\epsilon_t \in \mathcal{E}$ is an idiosyncratic shock, different among people. An individual does not have to stick to the same projects over time and she can choose portion $\phi_t \in [0, 1]$ of her savings to invest in high-risk high-return projects.

A financial institution provides two services to their customers in this simple model. First, a financial institution offers insurance for idiosyncratic shocks by pooling funds. Second, a financial institution selects projects when people apply for loans by inferring the true aggregate shocks, and tell them if they should stay in the relatively safe occupation or engage in the high-risk high-return business.[4]

Financial services, however, are not free. They require a one-time cost $q > 0$ to start using them and a per-period cost $(1 - \gamma) \in [0, 1]$ proportional to the investment amount. These costs can be thought

[4]An alternative interpretation is that the household puts money on deposit but then borrows to finance a project under advice from the bank. Average repayment is determined by either θ_t or δ. But if risky projects are undertaken, low ϵ_t households repay less, as if receiving insurance, and high ϵ_t households repay more, as if paying premia, so as to repay θ_t on average. That is, the debt repayment is allowed to vary with idiosyncratic shocks. Either way, we recognize that this specification of the financial sector's advantages may be extreme. One could imagine less-than-perfect risk-sharing, constrained by default or private information considerations, for example. One could also imagine less-than-perfect information about forthcoming realizations, as the number of bank clients engaged in any given activity (or sector/region) may be limited and, in any event, past experience in a given activity is only a limited guide to future shocks. However, this specification of financial services does make the model tractable.

of as combinations of intrinsic transactions costs and institutional impediments to a country's financial sector.[5]

In summary, those who are not using financial services accumulate assets according to

$$k_{t+1} = (\phi_t(\theta_t + \epsilon_t) + (1 - \phi_t)\delta)s_t, \tag{1}$$

and those who are using financial services accumulate according to

$$k_{t+1} = r(\theta_t)s_t \equiv \gamma \max\{\theta_t, \delta\}s_t. \tag{2}$$

We assume here formally that a financial institution has indeed a real informational advantage despite marginal costs, and that the risky asset is potentially profitable enough to attract positive investment, that is, the expected risky return dominates the safe return, even without advance information.[6]

Assumption 1.

$$E[r(\theta_t)] > E[\theta_t] > \delta > 0. \tag{3}$$

An individual chooses whether she uses financial service $d_t = 1$ or not $d_t = 0$, savings s_t, and portfolio share of risky projects ϕ_t to maximize her expected life-time utility:

$$E_1 \left[\sum_{t=1}^{\infty} \beta^{t-1} u(c_t) \right] \tag{4}$$

subject to the budget constraint

$$c_t = k_t - s_t - q\mathbf{1}_{d_t > d_{t-1}}, \tag{5}$$

where $\beta \in (0, 1)$ denotes the discount rate and $\mathbf{1}_{d_t > d_{t-1}}$ denotes an indicator function, which takes value 1 if an individual joins the financial system at t (i.e., $d_t > d_{t-1}$) and takes value 0 otherwise.[7] Though GJ restricts attention to the log contemporaneous utility, $u(c_t) = \log c_t$, we analyze as well the constant relative risk aversion (CRRA) utility function $u(c_t) = c_t^{1-\sigma}/(1 - \sigma)$ for $\sigma > 0$, where σ denotes the degree of relative risk aversion.

Because assets might accumulate unboundedly toward ∞ with a series of good shocks or deplete toward zero with a series of bad shocks, we introduce three assumptions that ensure the consumer's

---

[5]GJ and Townsend (1978) show this return is consistent with the return offered by a competing set of financial intermediaries. We do not pursue decentralized interpretation further. Of course, this specification of transactions costs begs for serious generalization, and extension to other kinds of costs, to distinguish among various possible kinds of involvement in the financial system. We can allow heterogeneous costs across different households/businesses, and indeed we do allow costs to vary with education and urban/rural status below.

[6]This is assumption C and part of assumption A of Greenwood and Jovanovic (1990).

[7]In practice, d_t will be zero for several periods and then jump to one and stay there, that is, no one will ever exit in this transitional growth model, and either $d_t > d_{t-1}$ or $d_t = d_{t-1}$. See below.

optimization problem (4) is well defined.[8] Note that these assumptions place restrictions on parameter values for the calibration exercise.

First, the cumulative distributions of θ_t and ϵ_t are assumed to be time invariant and denoted as $F(\theta_t)$ and $G(\epsilon_t)$, respectively, with compact supports:[9]

Assumption 2. *Let* $\Theta = [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_{++}$ *and* $F : \Theta \rightarrow [0, 1]$. *Let* $\mathcal{E} = [\underline{\epsilon}, \overline{\epsilon}] \subset \mathbb{R}$ *and* $G : \mathcal{E} \rightarrow [0, 1]$, *with* $E[\epsilon_t] = 0$.

We sometimes refer to total return $\eta_t \equiv \theta_t + \epsilon_t \in [\underline{\eta}, \overline{\eta}]$, and its cumulative distribution is denoted as $H : \Theta + \mathcal{E} \rightarrow [0, 1]$.

Second, the life time utility (4) must not explode. This is ensured by the following assumption, limiting the expected return, adjusted by risk aversion, to be smaller than $1/\beta$:[10]

Assumption 3. $\beta E[(r(\theta))^{1-\sigma}] < 1$.

Finally, as we focus on perpetual growth cases, it seems natural that the optimized life time utility have a real value bounded from below. A sufficient condition is to make the safe return sufficiently high, that is, greater than $1/\beta$,

Assumption 4. $\beta\delta > 1$.

B. Recursive Formulation

Because it is difficult to obtain analytical solutions that maximize the life time utility (4) for non-participants, we use numerical methods. More specifically, we use dynamic programming, transforming the original maximization problem at the initial date to a recursive maximization problem conditional on assets and participation status to the financial system in each period.[11] Following the notation of GJ, we define $V(k_t)$ as the value for those who have already joined financial intermediaries today, and $W(k_t)$ as the value for those who have not joined today but have

---

[8]Note that $u(\infty) = \infty$ for $\sigma \geq 1$ and $u(0) = -\infty$ for $\sigma \leq 1$. Three assumptions 2–4 together with some measurability requirements guarantee existence of the maximum and optimal decisions. Assumption 4 also makes the economy to grow perpetually. See the proofs in the working paper, Townsend and Ueda (2001).

[9]This assumption is also part of assumptions A and B of GJ.

[10]When $\sigma = 1$, assumption (3) becomes $\beta < 1$. Although assumption 3 applies to participants, $\beta E[\eta^{1-\sigma}] < 1$ is the analogue condition for nonparticipants by the same argument. This latter assumption is not necessary, because everyone eventually participates in the financial system, as is shown below.

[11]With some additional technical assumptions we can establish the equivalence of solutions between these two formulations. See proofs in the working paper, Townsend and Ueda (2001).

an opportunity to do so tomorrow. Also, we introduce a pseudo $W_0(k_t)$ as the value for those who are restricted to never joining. Explicit forms of these value functions are following:
for participants

$$V(k_t) = \max_{s_t} u(k_t - s_t) + \beta \int \max\{W(k_{t+1}), V(k_{t+1})\} dF(\theta_t) \tag{6}$$

subject to the wealth accumulation process (2);
for nonparticipants

$$W(k_t) = \max_{s_t, \phi_t} u(k_t - s_t) + \beta \int \max\{W(k_{t+1}), V(k_{t+1} - q)\} dH(\eta_t) \tag{7}$$

subject to the wealth accumulation process (1); and
for never-joiners

$$W_0(k_t) = \max_{s_t, \phi_t} u(k_t - s_t) + \beta \int W_0(k_{t+1}) dH(\eta_t) \tag{8}$$

subject to the same wealth accumulation process (1).

We can also establish, as in GJ, that participants will never terminate membership:[12]

$$W_0(k_t) \leq W(k_t) < V(k_t). \tag{9}$$

This implies that V is the only relevant branch on the right-hand-side of the functional equation (6).

Note that in the notation of GJ, the entering decision is made next period, not today. We can write an equivalent formulation in which the entering decision is made at the beginning of each period. We use this formulation to derive some analytical properties as well as to obtain numerical solutions. It is simply defined as

$$Z(k_t) \equiv \max_{d_t \in \{0,1\}} \{W(k_t), V(k_t - q)\}, \tag{10}$$

where $V(k_t - q)$ represents the value for *new* participants today. Then, the nonparticipant's value W in the GJ formulation can be also written as

$$W(k_t) = \max_{s_t, \phi_t} u(k_t - s_t) + \beta \int Z(k_{t+1}) dH(\eta_t). \tag{11}$$

C. Solutions of Value Functions and Policies

For non-participants with value $Z(k)$, the savings s and the portfolio share ϕ are functions of wealth k, and these need to be obtained numerically.[13] Since the economy grows perpetually, we cannot

---

[12]Again, see the proof in the working paper, Townsend and Ueda (2001).

[13]We omit time subscript t in the value functions because individuals face the same problem in each period given the current wealth level k. For detailed derivation of solutions in this section, see Townsend and Ueda (2001).

apply a standard numerical algorithm, which requires an upper bound and a lower bound of wealth level k. Fortunately, the participant's value $V(k)$ and the never-joiner's value $W_0(k)$ have closed form solutions together with the associated optimal savings rate and portfolio share. We utilize these two boundary value functions $V(k)$ and $W_0(k)$ to compute nonparticipant's values $W(k)$ and $Z(k)$. The numerical algorithm is described in the appendix.

Solution *V(k)*, the Participant's Value Function, and the Associated Policies

A participant's value function (6) is easily obtained under log utilities, by guessing and verifying, as in GJ.

$$V(k) = \frac{1}{1-\beta}\ln(1-\beta) + \frac{\beta}{(1-\beta)^2}\ln\beta + \frac{\beta}{(1-\beta)^2}\int \ln r(\theta)\, dF(\theta) + \frac{1}{1-\beta}\ln k. \tag{12}$$

The saving *rate* μ, defined as $\mu \equiv s/k$, total savings divided by beginning-of-period wealth, is equal to β for the log utility case.

More generally for CRRA utilities ($\sigma \neq 1$), we can also obtain the analytical formula for the value function $V(k)$ and the optimal saving rate μ^*:

$$V(k) = \frac{(1-\mu^*)^{-\sigma}}{1-\sigma}k^{1-\sigma}, \tag{13}$$

and

$$\mu^* = \{\beta E[r(\theta)^{1-\sigma}]\}^{1/\sigma}. \tag{14}$$

Solution *Wo(k)*, the Value Function for Those Never Allowed to Join the Bank, and the Associated Policies

Similarly, we can obtain analytical solutions for $W_0(k)$, the associated optimal saving rate μ^{**}, and the optimal portfolio share in the risky technology ϕ^{**}. For CRRA utility,

$$W_0(k) = \frac{(1-\mu^{**})^{-\sigma}}{1-\sigma}k^{1-\sigma}, \tag{15}$$

and

$$\mu^{**} = \{\beta E[e^{**}(\eta)^{1-\sigma}]\}^{1/\sigma}, \tag{16}$$

where $e^{**}(\eta) = \phi^{**}\eta + (1-\phi^{**})\delta$ is the optimized per unit return.[14]

---

[14]We can also show uniqueness of the optimal portfolio choice ϕ^{**}, which can take on the boundary values 0 or 1 (see Townsend and Ueda, 2001). Conditions for the boundary values are given by
(i) $\phi^{**} = 0$ if $E[\eta] < \delta$, that is, the safe return is sufficiently high (sufficient condition), and
(ii) $\phi^{**} = 1$ only if $E[1/\eta^{\sigma}] \leq 1/\beta\delta$, that is, the safe return is sufficiently low (necessary condition).

For the log utility, CRRA at $\sigma = 1$, the value function is

$$\begin{aligned} W_0(k) = &\frac{1}{1-\beta}\ln(1-\beta) + \frac{\beta}{(1-\beta)^2}\ln\beta \\ &+ \frac{\beta}{(1-\beta)^2}\int \ln e^{**}(\eta)\, dH(\eta) + \frac{1}{1-\beta}\ln k, \end{aligned} \tag{17}$$

and the optimal savings is $\mu^{**} = \beta$.

III. Analytical Characterization

A. Concavity of Transitional Value Functions

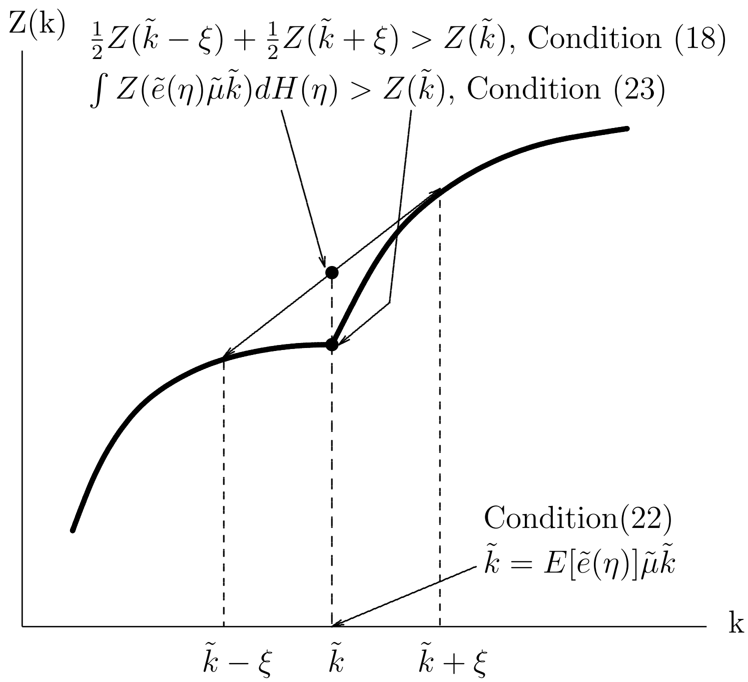
To simulate the economy, we need to make sure that we obtain optimal decisions by numerical methods. This would be difficult if there were multiple optimal decisions on savings, portfolio share, and entry to the financial system. This might be the case if the value function were not globally concave functions.[15] Indeed, the value functions for nonparticipants might not be concave, because the entry cost is a one-time fixed cost and this introduces a fundamental nonconvexity. Put differently, the value function $V(k)$ is strictly concave after entry and the value function $W(k)$ may be strictly concave before entry, but still the outer envelope, the value function $Z(k)$, which determines the entry point, might not be concave (see figure 1). If $Z(k)$ is not concave, $W(k)$ is no longer assured to be concave by definition (11).

However, with a concave period utility, an individual prefers to eliminate the nonconcave part of life time utility, if possible. Fortunately, the risky asset is a natural lottery which allows some convexification. Through the choice of her portfolio share in the risky asset, an individual can "control" the randomness of her capital in the next period and hence her lifetime utility from the next period on. If there remained a nonconcave part in the value function next period, increased variation in random returns from her investment would make her expected lifetime utility today bigger relative to nonrandom investment. Another way to see this result is to note that the outer envelope $Z(k)$ is supposed to reflect discounted expected utility for each given wealth k, but evidently higher values are possible, if the envelope is nonconcave, by a little more randomization in k today, something possible with greater stochastic investment yesterday. This need for randomization would imply that we have not found the optimal policy yet and that further iteration of the value function would be called for, so as to eliminate the nonconcave part eventually.

The result here is new to the literature on switching-state models, not only in the costly participation models of the financial system mentioned in the introduction, but also in labor search models. For example, in Danforth (1979), there may exist multiple solutions for an unemployed person to select his consumption level as well as when and which job offer he accepts over time. In a two-sided match model, Mortensen (1989) shows that multiple equilibria arise when the technology of

[15]Proof of concavity is at best implicit in GJ, and single-valuedness of the policy functions seems to be assumed implicitly in GJ though necessary for numerical computation here.

Figure 1. Nonconcave Function $Z(k)$



matching between workers and firms exhibits increasing returns to scale. Gomes and others (2001), a recent application of a labor search model to the business cycle literature, note that the concavity property for the life time utility with respect to asset holds for their simulation experiments, but they do not provide a proof or a sufficient condition. The core idea of our proof is that risky business serves as a natural lottery and that the optimal choice of risk makes the frontier of the life-time utility concave. Although the proof provided here is tailored to a specific model, the core idea behind the proof should make it possible to prove concavity of objective functions and uniqueness of policy functions in other switching-state models.[16]

We now show below that, under quite general assumptions, the optimal portfolio choice by an individual smooths out any nonconcave part and makes the overall value function concave. As a result, decisions (policy functions) become single-valued. Sufficient conditions to make this mechanism work are a reasonable degree of risk aversion and the existence of risk itself with some realistic properties as in assumptions 5, 6, and 7 below.

We start the proof with three basic properties of $Z(k)$: single-valuedness, upper semicontinuity, and monotonicity. The detailed proofs are in Townsend and Ueda (2001)[17] and here we simply outline the proofs. The life-time utility, which is the discounted sum of period utilities with costly switching of participation status over infinite periods, is shown to be upper semicontinuous. The Wierstrass theorem assures existence of the maximum as well as upper semicontinuity of the maximized life-time utility. The value function $Z(k)$ is shown to be equivalent to the maximized life-time utility

[16]We owe this point to a referee.

[17]We also show there another desired property, measurability of $Z(k)$.

given wealth k, implying single-valuedness and upper semicontinuity of $Z(k)$. Finally, given that a household can dispose of k freely, $Z(k)$ must be monotonically increasing.

With these three properties we establish a lemma: if $Z(k)$ were not a globally concave function, there should exist a region where a mid point of two $Z(k)$s evaluated at the edges of the region is greater than $Z(k)$ evaluated at the mid point of the region (see figure 1). Moreover, when $Z(k)$ is continuous in that region, a mid point of two $Z(k)$s evaluated at any symmetric pair of k inside the region is greater than or equal to $Z(k)$ evaluated at the mid point of the region. Formally, we write this as follows.

Lemma 1. *If $Z(k)$ were not a globally concave function, then there would exist an associated wealth level $\tilde{k}$ for any sufficiently small $\xi > 0$, such that*

$$\frac{1}{2}Z(\tilde{k}-\xi) + \frac{1}{2}Z(\tilde{k}+\xi) > Z(\tilde{k}). \tag{18}$$

In addition, if $Z(k)$ were also a continuous function inside the region $(\tilde{k}-\xi, \tilde{k}+\xi)$, then for any positive $\underline{\xi} < \xi$, the following weak inequality (19) holds:

$$\frac{1}{2}Z(\tilde{k}-\underline{\xi}) + \frac{1}{2}Z(\tilde{k}+\underline{\xi}) \geq Z(\tilde{k}). \tag{19}$$

See the proof in the appendix. Lemma 1 ensures that, when $Z(k)$ is a continuous function, a simple sum of any symmetric pairs within the range $[\tilde{k}-\xi, \tilde{k}+\xi]$ is larger than $Z(\tilde{k})$.[18] By taking a limit, integration of $Z(k)$ over the range should also be larger than $Z(\tilde{k})$, that is,

$$\int_{-\xi}^{\xi} Z(\tilde{k}+x)dx > Z(\tilde{k}). \tag{20}$$

In our model, however, integration is taken for the product of $Z(k)$ and $h(\eta)$, the combined probability density for idiosyncratic and aggregate shocks. To preserve an inequality similar to (20), it is sufficient for the probability density to exhibit positive mass almost everywhere and symmetric as in the following assumption.

Assumption 5. *Both of the cumulative density functions, $F(\theta)$ and $G(\epsilon)$, are non-degenerate and strictly increasing with full support. Moreover, their probability densities are symmetric.*

Assumption 5, together with Lemma 1, establishes that the integral of $Z(\tilde{k}+x)h(x)$ over a symmetric neighborhood would be larger than $Z(\tilde{k})$, when it is continuous. Intuitively, Lemma 1,

[18] For example, take $\underline{\xi} = \xi/2$, then (18) and (19) imply

$$\frac{1}{4}Z(\tilde{k}-\xi) + \frac{1}{4}Z\left(\tilde{k}-\frac{\xi}{2}\right) + \frac{1}{4}Z\left(\tilde{k}+\frac{\xi}{2}\right) + \frac{1}{4}Z(\tilde{k}+\xi) > Z(\tilde{k}).$$

which says a sum of each symmetric pair would be larger than or equal to the mid point, would be valid even $Z(k)$ were multiplied by $h(x)$, because $h(x)$ gives the same weight for each element of a pair. The details are given in the proof for lemma 2, but the same result can also be established even if $Z(k)$ were a discontinous function. In sum,

$$\int_{-\xi}^{\xi} Z(\tilde{k}+x)h(x)dx > Z(\tilde{k}). \tag{21}$$

We can map this relationship to our model for stochastic next period wealth. Formally, we define a hypothetical policy for savings $\tilde{\mu} \in (0,1)$ and portfolio share $\tilde{\phi} \in (0,1]$ at $\tilde{k}$ with the associated return $\tilde{e}(\eta) = \tilde{\phi}\eta + (1-\tilde{\phi})\delta$ that gives the average of capital next period equal to capital of this period, that is,

$$E[\tilde{e}(\eta)]\tilde{\mu}\tilde{k} = \tilde{k}. \tag{22}$$

Note that it is always the case that $E[\tilde{e}(\eta)] > 1$ and thus $\tilde{\mu} \in (0,1)$.

Lemma 2. *Lemma 1 and Assumption 5 assure that there exists a pair $\tilde{\mu}$ and $\tilde{\phi}$ such that integral over the value function $Z(k)$ over possible next period's capital starting from $\tilde{k}$ today is strictly larger than the value function evaluated at $\tilde{k}$ (see figure 1), that is,*

$$\int_{\underline{\eta}}^{\overline{\eta}} Z(\tilde{e}(\eta)\tilde{\mu}\tilde{k})dH(\eta) > Z(\tilde{k}). \tag{23}$$

See the proof in the appendix. Essentially, integration is a weighted average with the weight equal to the probability measure of shocks η, which is symmetric in the range $[\underline{\eta}, \overline{\eta}]$, and the portfolio choice ϕ determines the choice of the range of average on the next period capital $[\tilde{e}(\underline{\eta})\tilde{\mu}\tilde{k}, \tilde{e}(\overline{\eta})\tilde{\mu}\tilde{k}]$, which is analogous to $[\tilde{k}-\xi, \tilde{k}+\xi]$ in (21).

Two more assumptions are needed to prove concavity of $Z(k)$. One requires that households be sufficiently risk averse.

Assumption 6. *The period utility is at least weakly more concave than log utility, that is, $\sigma \geq 1$.*

The last assumption is a complex one, and we explain it below in detail.

Assumption 7. *The participant's savings rate μ^* satisfies*

$$\log(1-\mu^*) \geq \frac{1}{\sigma}\log(1-\beta) + \left(1-\frac{1}{\sigma}\right)\log\left(1-\frac{1}{E(\theta)}\right). \tag{24}$$

This assumption is sufficient to assure proposition 1, which follows below, that the value function $Z(k)$ is concave. For $\sigma = 1$ (log case), the right hand side of (24) is $\log(1-\beta)$, but the left hand side is also $\log(1-\beta)$ because $\mu^* = \beta$ in the log utility case. Hence the assumption 7 is always satisfied in the log utility case. The reason why the participant's savings rate μ^* appears in the condition for concavity of the nonparticipant's value function $Z(k)$ is that $Z(k)$ itself is too complex to analyze and we use the participant's value function $V(k)$ to obtain a stronger, sufficient condition. The participant's savings rate, nevertheless, is a good approximation of the average level of nonparticipant's saving rate and changes in the underlying parameters should have similar effects for both.

Assumption 7 is another restriction on the growth rate. It says that the savings rate cannot be too high, given the expected return from the risky asset $E(\theta)$. Since the savings rate is an endogenous variable, assumption 7 essentially pins down the parameter values such as the curvature of utility function. We can see this point clearly by simplifying assumption 7. Notice that the right hand side of (24) is just a linear combination of $\log(1-\beta)$ and $\log\left(1-\frac{1}{E(\theta)}\right)$ and, by assumptions 1 and 4, $1-\beta < 1-\frac{1}{\delta} < 1-\frac{1}{E(\theta)}$. Thus, with $\sigma \geq 1$ (assumption 6), the right hand side of (24) is at most $\log\left(1-\frac{1}{E(\theta)}\right)$. Hence the stronger but easy-to-check version of condition (24) is[19]

$$\mu^* = \left\{\beta E[r(\theta)^{1-\sigma}]\right\}^{1/\sigma} \leq \frac{1}{E(\theta)}. \tag{25}$$

Because the relative risk aversion σ increases, the stochastic savings becomes less attractive and the savings rate decreases (i.e., $\partial\mu^*/\partial\sigma < 0$). Hence, the curvature of the utility function should be sufficiently large, ceteris paribus, so that (25) is satisfied. Smaller discount rate β, larger γ (lower variable cost), and higher safe technology return δ work similarly.[20] Although assumption 5 reveals that variations in return are essential to ensure concavity for $Z(k)$, the effect of higher variation of θ on sufficient condition (25) is uncertain. Because $r(\theta) = \gamma \max\{\theta, \delta\}$, with $\sigma \geq 1$, is a convex and decreasing function of θ, higher variation of θ could make $E[r(\theta)^{1-\sigma}]$ higher (and thus the saving rate higher), but the left truncation of $r(\theta)$ at $\gamma\delta$ implies that higher variation of θ could make $E[r(\theta)^{1-\sigma}]$ lower.

The intuition is as follows: Through the choice of his portfolio share of the risky asset, an individual can control the randomness of his capital at the next period and his life time utility from the next period on. Note in figure 1 it is essential to have random returns to span the convex part of $Z(k)$. If there is a non-concave part in the value function, random returns from his investment may make his expected life time utility bigger than non-random investment in the neighborhood of non-concave part. This becomes true when his period utility is sufficiently concave. In figure 1, the wings of the "butterfly" are more salient, if both sides of the kink point are more concave. Finally, if randomization enlarged the frontier of the expected life time utility, further iteration of the value function would be called for so as to eliminate the non-concave part.

Finally, we have the needed proposition:

---

[19] Note that condition (25) is only a slightly stronger condition than assumption 3, which restricts the savings rate so that the life time utility is bounded by above. Indeed, with $\sigma \geq 1$, assumption 3 is always satisfied.

[20] Let $M = \beta E[r(\theta)^{1-\sigma}]$ so that $\mu^* = M^{1/\sigma}$. Then, $\frac{\partial\mu^*}{\partial\sigma} = \frac{\partial\mu^*}{\partial M}\frac{\partial M}{\partial\sigma}$ and $\frac{\partial\mu^*}{\partial M} = \frac{1}{\sigma}M^{(1-\sigma)/\sigma} > 0$. Hence, the sign of $\frac{\partial M}{\partial parameter}$ determines the sign of the effect on μ^* by a change in a parameter value. For example, $\frac{\partial M}{\partial\sigma} = -\beta E[r(\theta)^{1-\sigma}\log r(\theta)] < 0$ (note that $r(\theta) > 1$ for all θ) and $\frac{\partial M}{\partial\gamma} = (1-\sigma)E[r(\theta)^{-\sigma}\max(\theta,\delta)] < 0$ when $\sigma > 1$.

Proposition 1. *Under assumptions 5, 6, and 7, $Z(k)$ is concave.*

Proof. Under assumptions 5 and 6, we claim that, if assumption 7 (condition 24) holds, then $Z(k)$ is a concave function. We do not prove our claim directly, but prove its contrapositive: if $Z(k)$ is not a globally concave function, then the opposite of condition (24) holds.

Suppose $Z(k)$ were not a globally concave function, that is, $Z(k)$ does not have a convex subgraph.[21] Then, there must exist a point $\tilde{k}$ where integration over $Z(k)$ for possible realizations in the next period is outside the subgraph of $Z(k)$, as shown in lemma 2.

Consider now the value for the nonparticipant at the capital level $\tilde{k}$,

$$W(\tilde{k}) = \max_{\mu,\phi} u((1-\mu)\tilde{k}) + \beta \int Z(e(\eta)\mu\tilde{k})dH(\eta). \tag{26}$$

Since the policy $(\tilde{\mu}, \tilde{\phi})$ is possibly nonoptimal,

$$W(\tilde{k}) \geq u((1-\tilde{\mu})\tilde{k}) + \beta \int Z(\tilde{e}(\eta)\tilde{\mu}\tilde{k})dH(\eta). \tag{27}$$

By lemma 2 (condition 23) with assumption 5 and $\tilde{c} = (1-\tilde{\mu})\tilde{k}$,

$$W(\tilde{k}) > u(\tilde{c}) + \beta Z(\tilde{k}). \tag{28}$$

Here, we focus on the case $\sigma > 1$, but the following logic is applicable for the $\sigma = 1$ ($\log$) case.[22] By the closed solution of $V(k)$ as in equation (13),

$$\begin{aligned} u(\tilde{c}) =& \frac{\tilde{c}^{1-\sigma}}{1-\sigma} \\ =& (1-\beta)(1-\beta)^{-1}(1-\mu^*)^{\sigma}(1-\mu^*)^{-\sigma}\frac{\tilde{c}^{1-\sigma}}{1-\sigma} \\ =& (1-\beta)\frac{(1-\mu^*)^{-\sigma}}{1-\sigma}\left((1-\beta)^{\frac{-1}{1-\sigma}}(1-\mu^*)^{\frac{\sigma}{1-\sigma}}\tilde{c}\right)^{1-\sigma} \\ =& (1-\beta)V\left((1-\beta)^{\frac{-1}{1-\sigma}}(1-\mu^*)^{\frac{\sigma}{1-\sigma}}\tilde{c}\right). \end{aligned} \tag{29}$$

By substituting this for $u(\tilde{c})$ into the right hand side of inequality (28), and using $Z(\tilde{k}) = W(\tilde{k})$,[23] we obtain

$$W(\tilde{k}) > (1-\beta)V\left((1-\beta)^{\frac{-1}{1-\sigma}}(1-\mu^*)^{\frac{\sigma}{1-\sigma}}(1-\tilde{\mu})\tilde{k}\right) + \beta W(\tilde{k}). \tag{30}$$

---

[21] The subgraph of $Z(k)$ is the set $\{(k,z) \in \mathbb{R}_+ \times \overline{\mathbb{R}} : z \leq Z(k)\}$.

[22] We can take $\sigma \to 1$.

[23] $Z(k)$ is by definition equal to either $V(k-q)$ or $W(k)$, but $V(k-q)$ is concave. Hence, it must be equal to $W(k)$ at $\tilde{k}$, where $Z(k)$ is assumed to be non-concave.

Equivalently, using $\bar{k} \equiv (1-\beta)^{\frac{-1}{1-\sigma}}(1-\mu^*)^{\frac{\sigma}{1-\sigma}}(1-\tilde{\mu})\tilde{k}$,

$$W(\tilde{k}) > (1-\beta)V(\bar{k}) + \beta W(\tilde{k}). \tag{31}$$

We can simplify (31) to

$$W(\tilde{k}) > V(\bar{k}). \tag{32}$$

Note that, if $Z(k)$ is not globally concave, inequality (32) must hold.

Since $V(k) > W(k)$ for any k, as shown in (9), if inequality (32) holds, then it must be the case that

$$V(\tilde{k}) > V(\bar{k}). \tag{33}$$

Because of the definition of $\bar{k}$ and $V(k)$ defined in (13),

$$\begin{aligned} V(\bar{k}) &= V\left((1-\beta)^{\frac{-1}{1-\sigma}}(1-\mu^*)^{\frac{\sigma}{1-\sigma}}(1-\tilde{\mu})\tilde{k}\right) \\ &= (1-\beta)^{-1}(1-\mu^*)^{\sigma}(1-\tilde{\mu})^{1-\sigma}V(\tilde{k}), \end{aligned} \tag{34}$$

inequality (33) is equivalent to

$$1 > (1-\beta)^{-1}(1-\mu^*)^{\sigma}(1-\tilde{\mu})^{1-\sigma}. \tag{35}$$

But $\tilde{\mu} = 1/E[\tilde{e}(\eta)] \geq 1/E(\theta)$ and $\sigma \geq 1$ (assumption 6) implies $(1-\tilde{\mu})^{1-\sigma} \geq (1-1/E(\theta))^{1-\sigma}$. Hence, if inequality (35) holds, it must be the case that

$$1 > (1-\beta)^{-1}(1-\mu^*)^{\sigma}\left(1-\frac{1}{E(\theta)}\right)^{1-\sigma}. \tag{36}$$

By taking logarithms, inequality (36) is equivalent to

$$\log(1-\mu^*) < \frac{1}{\sigma}\log(1-\beta) + \left(1-\frac{1}{\sigma}\right)\log\left(1-\frac{1}{E(\theta)}\right), \tag{37}$$

which is the opposite to assumption 7 (condition 24). Therefore, under assumptions 5 and 6, assumption 7 implies that $Z(k)$ is a globally concave function.

$\square$

Here are straightforward implications of the proposition 1 (see the proof in the appendix).

Corollary 1. *$Z(k)$ is strictly concave. Given k_{t-1}, d_{t-1} and d_t, the optimal policy (μ_t, ϕ_t) is single-valued, and $\mu_t(k)$ and $\phi_t(k)$ are continuous functions on $\mathbb{R}_{++}$.*

Remark. [24] (i) It is an immediate result that $W(k)$ satisfies the same properties as $Z(k)$ described in proposition 1 and corollary 1.
(ii) Using proposition 1, we can show that $Z(k)$ is continuous and differentiable on $\mathbb{R}_{++}$.

[24] See Townsend and Ueda (2001) for the proof.

B. Decisions in Transition and in the Long-run

As savings and portfolio decisions are single-valued for nonparticipants, we can now characterize their properties. First, because nonparticipants prepare to pay the future fixed entry fee, their saving rate will be higher than participants' μ^*, that is, for nonparticipants,

$$\mu(k) > \mu^*. \tag{38}$$

Second, although it will be shown below that everyone eventually joins the financial system, those who have very little wealth act as if they would never be able to join a bank. In other words, very poor people have almost the same value and policies (μ^{**}, ϕ^{**}) as those who never have the opportunity to join a bank. In the short run, agents with very small capital spend little effort to accumulate capital to join the bank, that is, for the very poor, for all $\epsilon > 0$, there exists some k_ϵ sufficiently small such that

$$\sup_{k \in (0, k_\epsilon]} \left| \frac{s(k)}{k} - \mu^{**} \right| < \epsilon, \tag{39}$$

and

$$\sup_{k \in (0, k_\epsilon]} |\phi(k) - \phi^{**}| < \epsilon. \tag{40}$$

These are extensions of proposition 2 and 3 of Greenwood and Jovanovic (1990) to the CRRA utility case, but the proof is the same and we omit it.

In the long run, every household becomes rich as its wealth grows unboundedly, implying that no one actually takes the value of W_0. Namely, we can show[25] the following properties:
(i) Participants in the bank almost surely accumulate their wealth to exceed any $K < \infty$ in the long run as $t \to \infty$;
(ii) Those who were never allowed to join the bank would almost surely accumulate their wealth to exceed any $K < \infty$ in the long run as $t \to \infty$; and
(iii) Everyone eventually participates in the bank, almost surely, and hence $W(k) > W_0(k)$ for all k.

Property (ii) implies that even a household outside the financial system and with only a little wealth, eventually accumulates wealth, though stochastically, to a level greater than or equal to some critical capital level k^*. At that point, it will join the financial system and hence property (iii) must be true. As inequality (9) suggests, it will stay inside forever, that is, $d_{t+1} = 1$ if $d_t = 1$. Property (i) then implies that its wealth grows unboundedly.

We do assume that a household joins the financial system whenever it is indifferent,[26] that is, $d_t(k) = 1$ when $W(k) = V(k - q)$. In sum, we let k^* denote the *critical capital level*, that is, the least capital level at which the values of $W(k)$ and $V(k - q)$ coincide. This level is, of course, uniquely determined. Then, the participation decision for those whose initial asset level is less than

[25] See the proofs in Townsend and Ueda (2001).

[26] As in the standard competitive theory of firm, firms are considered to operate under zero profit.

the critical capital level becomes single-valued on the equilibrium growth path and thus d_t is monotonically (weakly) increasing in wealth. We use this property for simulation.[27]

C. Population Distributions

Although each household's return is not affected by others' choices, it does depend on its wealth. As a consequence, "macroeconomic" variables such as the growth rate of per capita income, participation rate, and overall inequality measures vary with the entire wealth distribution of participants and nonparticipants.

We can define the cumulative transition function for nonparticipants from period t capital k to period $t+1$ capital k' as follows

$$\Psi(k';k) \equiv prob[k_{t+1} \leq k' | k_t = k]. \tag{41}$$

In theory, we construct $\Psi(k';k)$ as defined in (41) from the law of motion of wealth (1), given the optimal policies (d_t, s_t, ϕ_t) and the distributions of shocks $F(\theta)$ and $G(\epsilon)$. In practice, we compute the optimal policies (d_t, s_t, ϕ_t), each as functions of k, and then we approximate $\Psi(k';k)$ on a discrete wealth grid using the numerical policies and assumed distribution functions $F(\theta)$ and $G(\epsilon)$ with a continuum of agents.

Assuming nonparticipant's initial wealth is below critical capital level k^*, we can recursively derive the wealth distribution at each period $t \geq 1$ for nonparticipants. To derive the wealth distribution next period, we exclude the population who accumulate wealth more than k^*, because at k^* people join the financial system. We need to distinguish nonparticipants in the last period into two types based on status in this period, either new entrants or continued nonparticipants.

Define $M_{t-1}(k)$ as the size of the population (cumulative distribution) in period $t-1$, who were outside of the intermediated sector at the last period $t-1$ and have a capital stock $k_t \leq k$ in the current period t. Note that $M_{t-1}(k)$ includes new entrants, who accumulate more than k^* in this period. We therefore construct the cumulative distribution $M_t(k')$ of wealth in the next period, for nonparticipants in this period, by integrating the transition function with respect to this period's capital over the restricted range $[0, k^*]$ for nonparticipants population in the last period,

$$M_t(k') = \int_0^{k^*} \Psi(k';k) dM_{t-1}(k). \tag{42}$$

[27] The remaining theoretical question is whether the participation decision d_t is single-valued in an off-equilibrium path, where a nonparticipant's initial wealth is larger than the critical capital level. Single-valuedness and monotonicity of the participation decision for all levels of wealth could be shown. However, for our empirical (simulation) purposes, we do not need this, because we make nonparticipants join the bank when their wealth first reaches the critical capital level and because we assign initial participation status to people at $t = 0$ depending on whether they have wealth above the critical capital level.

In a similar manner using (2), we can define $\hat{M}_t(k')$ as the population distribution of the participants at t who have a capital stock $k_{t+1} \leq k'$ in period $t+1$ and $\hat{M}_{t-1}(k)$ for participants at $t-1$ who have wealth $k_t \leq k$ in period t. In each period, we add new entrants, who were nonparticipants in the last period $t-1$ but accumulate wealth more than k^*, to the pool of participants. These new entrants are likely to have net wealth less than k^* after they join the financial system because they pay the fixed cost q to do so.

Now with these cumulative wealth distributions, we can find the ex post gross growth rate of capital k_{t+1}/k_t from t to $t+1$, namely (43) below for those who are outside of financial system both in period $t-1$ and t,

$$g_w(k_t, \theta_t, \epsilon_t) \equiv \{\phi(k_t)(\theta_t + \epsilon_t) + (1 - \phi(k_t))\delta\}\mu(k_t), \tag{43}$$

(44) below for participants both in period $t-1$ and t,

$$g_p(k_t, \theta_t) \equiv r(\theta_t)\mu^*, \tag{44}$$

and (45) below for new participants[28] who were outside in period $t-1$ but join the bank at t,

$$g_n(k_t, \theta_t) \equiv \frac{r(\theta_t)(k_t - q)}{k_t}\mu^*. \tag{45}$$

The economy-wide gross growth rate is easily obtained as the evolution of the population average wealth level. Specifically, we need an expression for the average wealth level of the economy in the current period K_t and in the next period K_{t+1}. K_t is defined as:

$$K_t \equiv \int_0^\infty k dM_{t-1}(k) + \int_0^\infty k d\hat{M}_{t-1}(k). \tag{46}$$

From this, the economy-wide gross growth rate is calculated essentially as K_{t+1}/K_t. That is, the economy-wide growth rate is basically a wealth-weighted average of the growth rate of each household, so it reflects rich households' growth rates more than poor households'. This is clear if we write K_{t+1} in terms of period t's population distributions M_{t-1} and $\hat{M}_{t-1}$, as below:[29]

$$\begin{aligned} E[K_{t+1}] = \int_{\underline{\epsilon}}^{\overline{\epsilon}} \int_{\underline{\theta}}^{\overline{\theta}} \Big[\int_0^{k^*} g_w(k, \theta, \epsilon) k dM_{t-1}(k) + \int_{k^*}^\infty g_n(k, \theta) k dM_{t-1}(k) \\ + \int_0^\infty g_p(k, \theta) k d\hat{M}_{t-1}(k) \Big] dF(\theta) dG(\epsilon). \end{aligned} \tag{47}$$

In sum, the expected economy-wide gross growth rate $E[K_{t+1}]/K_t$ in transition, denoted by x_{Gt}, is derived from formulas (46) and (47). As is clear from (46) and (47), it is a function (denoted Q

---

[28]Note that the new participants must pay fixed cost q before they save.

[29]To distinguish each integral, we write the domains of integrals explicitly here.

below) of $M_{t-1}(k)$ and $\hat{M}_{t-1}(k)$, the entire wealth distribution of participants and nonparticipants, respectively:

$$\begin{aligned} x_{Gt} &= Q(M_{t-1}(k), \hat{M}_{t-1}(k)) \\ &\equiv \frac{E[K_{t+1}]}{K_t}. \end{aligned} \tag{48}$$

As $t \to \infty$, x_{Gt} approaches $\mu^* E[r(\theta_t)]$, the long-run, steady-state growth rate in which (almost) everyone participates in the financial system.

Apparently, the expected participation rate at t, denoted by x_{Pt}, can be also written as a function (P) of the entire wealth distribution of participants and nonparticipants:

$$\begin{aligned} x_{Pt} &= P(M_{t-1}(k), \hat{M}_{t-1}(k)) \\ &\equiv \frac{\int_0^\infty dM_{t-1}(k)}{\int_0^\infty dM_{t-1}(k) + \int_0^\infty d\hat{M}_{t-1}(k)}. \end{aligned} \tag{49}$$

We use the Theil index as a measure of inequality. Unlike the Gini coefficient, it is decomposable into subpopulations, such as participants and nonparticipants. This is a nice property that simplifies analytical and numerical calculations.[30] The expected Theil index x_{It} can be written also as a function (I) of the entire wealth distribution of participants and nonparticipants:

$$\begin{aligned} x_{It} &= I(M_{t-1}(k), \hat{M}_{t-1}(k)) \\ &\equiv \int_0^\infty \frac{k}{K_t} \log\left(\frac{k}{K_t}\right) dM_{t-1}(k) + \int_0^\infty \frac{k}{K_t} \log\left(\frac{k}{K_t}\right) d\hat{M}_{t-1}(k). \end{aligned} \tag{50}$$

In practice, we need to approximate these "macro" variables $(x_{Gt}, x_{Pt}, x_{It})_{t=1}^\infty$ numerically. We compute the optimal savings rate $\mu(k)$ and the portfolio share $\phi(k)$ for nonparticipants, and construct the transition function (e.g., (41) for nonparticipants) over a set of discrete wealth levels as an approximation to the continuum space of wealth. We also have closed-form policy solutions for participants. We preserve the assumption of a continuum households, as noted in the discussion above, just after the introduction of the transition function $\Psi(k'; k)$ in (41). We refer to the sequence $(x_{Gt}, x_{Pt}, x_{It})_{t=1}^\infty$ as the expected path in the sense of integration over the population distribution of idiosyncratic shocks ϵ_t and expectation with respect to the distribution of aggregate shocks θ_t.

It is clear that the macro variables are functions of entire wealth distributions of participants and nonparticipants. These distributions evolve over time and reach the steady state growth path only after everyone joins the financial system. Until then, we cannot apply econometrics based on the assumption that these macro variables are stationary series, even after taking logs and lags, as clear in the formulae (48)–(50).

[30]See, for example, Cowel (1995) and Deaton (1997) for further properties of general entropy class measures of inequality. Our version here is a limit form of the Theil index defined for finite populations.

IV. Thai Economy: Growth, Inequality, and Financial Deepening

The economy to which we take the model is the Thai economy, concentrating on its emerging market growth phase, 1976–96, prior to the financial crisis of 1997. Needless to say, we do not attempt in this paper to analyze the crisis itself. Rather, we concentrate on the prior growth period, to see if we can understand this through the lens of the model, as a financial transition.[31]

We use various, multiple sources of data for calibration, estimation, and general discussion of the Thai economy. The national income accounts are constructed by the National Economic and Social Development Board, the NESDB. Credit and monetary aggregates as well as savings are provided by the Bank of Thailand. A complete village census with interviews of headmen was conducted by the Community Development Department, the CDD, biannually starting in 1986. A nationally representative *Socio-Economic Survey* covering income and expenditures, the SES, has been implemented at a substantial scale starting in 1976 with over 11,000 households, then repeated in 1981, and finally biannually after 1984.[32] In addition, we draw on the Townsend-Thai data,[33] a specialty cross-sectional May 1997 survey of 2,880 households in the central and northeastern regions, with measures of income, wealth, financial sector participation, savings and credit, and other items.

The Thai economy displays growth, financial deepening, and increasing (and eventually decreasing) inequality. Before the financial crisis of 1997, the Thai economy had grown rapidly. The NESDB numbers show growth in gross domestic product, ranging from 3 to 7 percent from 1976 to 1986, then a relatively high and sustained average growth of more than 8 percent for 1986–94, especially high in the 1987–89 period, and finally tailing off somewhat to 4 percent by 1996. The per capita income numbers from the SES are similar, with an overall average at 5 percent, relatively low at 2 percent for 1976–86, then high at 9 percent for 1986–92, and lower but still high at 7 percent for 1992–96.

Overall financial deepening is apparent in the macro aggregates. The ratio of M2/GDP rises steadily, surpassing the U.S. by 1992. Similar movements are apparent in M3/GDP, rising even faster in the 1988–94 period and also total credit/GDP, as reported in Klinhowhan (1999). Total credit extended by commercial banks increased with a particular surge in credit received by firms in the 1986–90 period. Likewise, restricting attention to rural areas, credit from the Bank for Agriculture and Agricultural Cooperatives (BAAC, a rural development bank) as a percentage of agricultural output increased throughout the 1981–97 period.

At a more micro level, from the CDD data, the fraction of village headmen reporting access to commercial bank credit rises from 0.26 to 0.41 in the 1986 to 1994 period, and to the BAAC from

---

[31]We also take the view that to understand the crisis and the subsequent recovery one must understand the growth that preceded it.

[32]We interpolate the bank access and inequality variables linearly for missing years.

[33]Detailed information is available in Townsend and others (1997), and also at the web page: http://www.src.uchicago.edu/users/robt.

0.80 to 0.92. At the level of households in the *Socio-Economic Survey*, respondents are asked whether they had changes in assets and/or liabilities from various financial institutions due to a transaction by any member in the previous month. Though no doubt noisy and off in levels, this measures rises from 6 to 26 percent, 1976–96. This is the measure of changing participation we use below. Further, we can stratify by occupation, education, and urban/rural status and see the same upward trend on this extensive margin for all groups. Though the access numbers are higher for those in urban areas and those with secondary or higher education, the expansion is particularly evident for villagers, from 5 to 23 percent, and those with primary education, from 5 percent to 22 percent.

We also emphasize the apparent effect of wealth constraints within each of these categories. We use the SES data to extract a latent-index measure of wealth and thus plot estimated wealth against that same SES measure of financial participation. Specifically, as in Jeong (2000), the SES provides ownership information on twenty household assets, and the latent variable is constructed to try to best explain the cross sectional variance. This is given a crude value in Thai baht by multiplying against the rental value of the respondent's house, though one should not take the assigned values literally (and we rescale below relative to the transactions cost). Then average access (again, transaction in the previous month) can be plotted by wealth deciles. In every survey year, even as late as 1996, these profiles are distinctly upward sloping with somewhat higher slopes at relatively high wealth levels. The theory with no heterogeneity in costs would predict a 0–1 jump up at some threshold wealth level but mitigated locally by the incurred fixed costs. We take these figures as prima facie evidence of wealth constrained choice of access to the financial system, as the theory would suggest. Of course the theory is giving us the structure to interpret the data. Causality cannot be inferred from these data alone.

As the model would suggest, beneath the growth of income and financial deepening lies relatively high and increasing inequality. The Gini measure of income inequality computed from the SES rises slowly but steadily from 0.42 in 1976 to 0.54 in 1992, then falls in the end to 0.50 by 1996. This is high for Asia, but lower than in some Latin American countries. This inequality reflects disparities in regional and rural/urban growth rates and appears to be related to factors discussed in various literatures, in particular, as we emphasize here, wealth-constrained access to the formal financial system.

Logically, we now focus on the financial transition. We formally view the Thai economy through the lens of the model to see if we can interpret some of these data and document anomalies.

V. Setup for Numerical Analysis

In the following sections, we analyze quantitative properties of the model by looking at numerically constructed expected paths, as described in section III. C, as well as Monte-Carlo simulations. Specifically, we examine the model in three ways: the expected path, the best-fit simulation, and the confidence region. With the expected path, we are able to see how well the model replicates the actual economy on average. With the best-fit simulation, we are able to see how the model can replicate the actual Thai data if we pick a particular path under a covariance-normalized mean squared error metric. With the confidence region, we are able to see whether the actual data lies within standard error bands using a test statistic. Before proceeding with these analyses, we pick reasonable values for the parameters, as in calibration exercises of the business cycle literature, then compute the value and policy functions using numerical methods.[34]

A. Setting Parameters

First, we set preference parameters following the business cycle literature.[35] Namely, we set the value of the discount rate $\beta = 0.96$. We report the log utility case primarily, but in a later section we also report $\sigma = 1.5$ as a robustness check.

Second, we set the technology parameters using Townsend Thai data (Townsend and others 1997). We use income-to-capital ratios to estimate the technology parameters for those not in the financial system. The survey shows that the median net return from capital investment in subsistence agriculture, which we regard here as a crude approximation to the safe project δ, at 5.4 percent in 1997. For the idiosyncratic shocks to the risky project, we also use income-to-capital ratios, but for those in nonagricultural business with no access to the financial sytem, and set the support of idiosyncratic shocks ϵ as $\mathcal{E} = [-0.6, 0.6]$. This is the range of returns or income-to-capital ratios from the bottom 1 percent to the top 99 percent.[36]

Pinning down the parameters on aggregate shocks turned out to be somewhat difficult. We know that the difference between the minimum and maximum real per capita growth rate from 1976 to 1996 is about 8.7 percent, and according to the model with projects selected by the financial sector, underlying variation of the aggregate shocks would be yet larger. Thus we assume the *range* for the aggregate shocks θ at 10 percent.[37] We vary the *mean* of θ and pick the support of θ as $\Theta = [1.047, 1.147]$ to minimize sum of squared errors of the actual GDP growth rate and the model prediction under some additional assumptions.[38] This is the only part of calibration which uses dynamic data. This is our benchmark.

---

[34] The numerical algorithm is described in detail in Appendix II.

[35] See, for example, Kydland and Prescott (1982).

[36] The more extreme tails no doubt contain noisy outliers.

[37] If our model were the true underlying mechanism, the observed aggregate GDP growth rate (and TFP, if it could be calculated) would not be a process generated by a stationary and ergodic process. Thus, the mean and variance of the GDP growth rate would not represent those of the underlying aggregate shock.

[38] In this search for the mean of θ, we cannot, for computational reasons, continuously update the optimal policy functions of the model. We thus make several assumptions and approximate the predictions of the model as we search for a plausible value for the mean of θ. Specifically, we assume the following: the actual participation rate is as observed in the data; savings rate of non-participants is 0.4 percent higher than participants regardless of wealth level; risky asset holding by non-participants are 30 percent of their savings; and households lose 1/3 of their assets upon their entry to the financial system. These are the typical values we observe in various parameter values, but it is not guaranteed that the optimal policies under the distribution of shock coincides with these simplistic assumptions.

Surprisingly, this benchmark makes the model's expected paths of financial deepening and inequality match with the actual Thai data on average over time. However, the predicted growth rate is about 2 percent lower than the actual one, and thus we also report an alternative case with 2 percent higher aggregate shocks, that is, $\Theta = [1.067, 1.167]$. This indeed makes the model's expected path of wealth growth match with the actual Thai GDP growth on average over time (see below).

The fixed cost q is a free parameter, and we take it to be $q = 5$ in model units of capital. By comparing the critical capital level k^* in the model units and k^* in the actual data in Thai baht, we find a scalar or "exchange rate" between the model units and the actual Thai baht. The critical capital level in model units is obtained by computing the value functions, namely $k^* = 15$ under the benchmark parameter values.[39] The critical capital level in the actual data is estimated using the SES and the observed fraction participating in 1976. That is, we use the wealth distribution of 1976 from SES of Thailand as the initial condition[40] (in 1990 baht), following Jeong (2000). We also use the information about participation in the financial system from the same SES. According to that survey, the fraction of the population who had access to the financial system was 6 percent in 1976. The estimated cumulative distribution of wealth in 1976 shows that people who had wealth of more than 220,000 baht[41] were 6 percent of the population in 1976. Since the critical level of the model is $k^* = 15$, we set the scalar or "exchange rate" as about 15,000 baht per model unit capital (in 1990 baht) to generate 6 percent participation in 1976.[42]

These parameters must satisfy the assumptions described in the model section. For example, we made assumption 1, $E[r(\theta)] > E[\theta] > \delta$. This condition implies that the variable cost $1 - \gamma$ cannot be large, and we assume a zero variable cost, $1 - \gamma = 0$ for the benchmark case but run some robustness checks for other values (e.g., 2 percent). Also, we would like to see some variation in the function $\phi(k)$, implying an interior solution of ϕ^{**} for $W_0(k)$ is desirable. This is checked numerically, but essentially, the mean risky return $E[\theta]$ with some adjustment for risk aversion cannot be much larger than the safe return δ.

All these numbers are summarized in table 1.

[39]The fixed cost turns out one third of wealth at the entry point. This might seem large, but if we take into account construction of bank branches and roads, it is not obviously too large.

[40]We include all 10,619 households after dropping about 600 households that report no income, and take sample weights into account. See Jeong (2000) for further discussion of the estimation of wealth.

[41]It is evaluated at the 1990 price level.

[42]In the simulations below, we need to approximate the initial wealth distribution. Assuming no new entrants in 1976, we divide the entire wealth distribution at $k^* = 220,000$ baht into two parts, one below k^* as the nonparticipants' wealth distribution $M_0(k)$ and the other as the participants' $\hat{M}_0(k)$. Also note that the "exchange rate" possibly varies with parameter values, because different parameter values produce different k^* in the model units but corresponding values in the data are always the same, $k^* = 220,000$ baht (except for a numerical approximation error).

Table 1. Parameter Values

	σ	q	δ	θ	ϵ	β	γ
$Benchmark$	1	5	1.054	$[\mathbf{1.047}, \mathbf{1.147}]$	$[-0.6, 0.6]$	0.96	0
$Alternative$	1	5	1.054	$[\mathbf{1.067}, \mathbf{1.167}]$	$[-0.6, 0.6]$	0.96	0

B. Value and Policy Functions

Computed value functions for the benchmark parameter values are shown in figure 2. $W(k)$ is always between $V(k)$ and $W_0(k)$ as discussed in section II.B. It approaches $W_0(k)$ as k goes to zero, and approaches $V(k-q)$ as k goes to ∞, as discussed in section III.B. The critical level of capital to join the bank is $k^* = 15$, where $W(k)$ and $V(k-q)$ cross.

The saving rate of nonparticipants increases with their wealth level up to near the critical level of capital that determines the entry decision, and then decreases slightly (see figure 3).[43] This is due to intertemporal consumption smoothing, preparing for the payment of the fixed fee. The portfolio share of risky assets varies in figure 3 as expected around the optimal level ϕ^{**} under $W_0(k)$, the value function of those who never enter the bank. It increases first and then decreases. It is, however, almost always *larger* than ϕ^{**} for $k < k^*$. As anticipated, nonparticipants put their wealth in the risky asset as a natural lottery to convexify their life-time utility (value function). For small levels of capital, the figures show that the saving rate and portfolio share approach those of those who never join the bank. This illustrates the analytical results (39) and (40).

These value and policy functions are very similar for the alternative parameter values $\Theta = [1.067, 1.167]$. One of the differences is the critical capital level to join the financial system k^*, which is smaller. Also, the portfolio share of risky assets is higher. Those results are natural consequences of the higher return (but the same risk) of risky assets.

Using these numerically obtained savings and portfolio share functions, we simulate the economy, given the initial wealth distribution. Figure 4 shows the simulated wealth evolution over 30 periods under the benchmark parameter values. We will discuss further how to evaluate the model with the actual Thai data, but, before proceeding, we would like to make it clear that simple regression studies are not well suited for a study of transitional economic dynamics.

VI. Growth, Inequality, and Financial Deepening: Spurious Regressions

We will show here that regression studies cannot capture true linkages among economic growth, financial deepening, and inequality, when the underlying economy is on a transitional growth path. Namely, we ask here whether our canonical model can generate data consistent with the findings in

[43] Note that the graph contains the counterfactual nonparticipants' policy over the capital range more than the critical value $k^* = 15$.

the empirical literature. Overall, the regression results are favorable: effects are found as in King and Levin (1993) and Forbes (2000). However, as we might expect for a transitional economy, we find these effects to be unstable, and this is also in line with some recent empirical studies. For example, Banerjee-Duflo (2000) reports on unstable and nonlinear relationship between inequality and growth and Favara (2003) reports similar findings between financial depth and growth.

Specifically, we conduct an experiment: we fix the benchmark economy, populate it with 1002 households respecting the initial 1976 Thai wealth distribution, and then draw idiosyncratic shocks in the population and aggregate temporal shocks for 30 years. We do this experiment 1000 times, with different shocks, generating in effect panel data for 1000 (artificial) countries. We then revisit these countries after 1976 to examine their status in later years.[44]

The advantage of any formal, structural model of growth is that the mechanism or "drivers" are made clear. Here for example, given common initial inequality in the wealth distribution and the parameters of technology and preferences, the drivers are the realized draws of idiosyncratic and aggregate shocks. There are stationary aspects to the model: household savings and portfolio decisions at date t, and hence the likelihood of financial participation at date $t+1$, are all determined by current wealth and current participation status. But aggregate growth, inequality, and overall financial deepening are not stationary time series even after taking logs and lags. They are all endogenous and all determined by these underlying shocks and decisions in complex and nonlinear ways.[45] Note that these complex dynamics are not only found in our canonical model, but also in many other theoretical models that depict endogenous financial deepening, inequality, and growth.

King and Levine (1993) report that there is a robust positive relationship between "initial" 1960 financial depth and subsequent growth, averaged over 1960 to 1989. They conclude that financial services stimulated growth. Here we regress 20 year average growth rates on the "initial" 1985, or 1980, level of financial depth, controlling for the initial log level of GDP (as created by 5 or 10 years of early model history).[46] Likewise Forbes (2000) replicates a typical finding in the empirical literature: a robust negative relationship between "initial" inequality in 1965 and average growth from 1965 to 1990. Here we regress 20 year average growth rates on initial 1985, or 1980, levels of

---

[44]Here, we use the Gini coefficient as a measure of inequality to be consistent with the empirical literature.

[45]If the data came from a stochastic steady state, there are some ways to normalize nonstationary time series so as to be represented as stationary time series. But, in this paper, the macro data are by construction taken from transient states. From the equation (47) of evolution of aggregate capital, the macro variables are functions of the underlying wealth distributions of participants $M_t(k)$ and nonparticipants $\hat{M}_t(k)$. These distributions change forms over time. Before reaching the steady state, they are transient, as shown in figure 4, never coming back to have the same shape, even after any macro level normalization is taken, due to the nonlinear savings and portfolio functions of individuals. Therefore, the stationarity assumption, that error terms of simple regressions are drawn from the same distribution over time or over countries, is not valid here. Moreover, the ergodic assumption, that error terms of simple regressions over time or over countries can reveal the underlying error distributions with a large sample, is not valid here either.

[46]The GDP level is normalized to one at 1976 in the regressions.

Table 2. Spurious Growth Regression Results: Long-Run Effects of Initial Financial Depth and Inequality onto Growth

Estimation Method	(1)	1985 as initial period (2)	(3)	(4)	1980 as initial period (5)	(6)
Constant	3.5110	7.0600	12.7343	3.6189	8.3826	11.1520
	(11.8937)	(4.7252)	(11.7274)	(10.8625)	(3.9869)	(6.5943)
Initial Financial Depth	4.7298		−36.0538	3.7519		−55.9250
	(1.7914)		(−31.0114)	(0.9869)		(−26.0722)
Initial Inequality		−5.2447	−8.8110		−5.8565	−1.9792
		(−1.8188)	(−4.2517)		(−1.4235)	(−0.6003)
Initial GDP	0.4324	0.4284	4.2420	0.5488	1.8229	4.4722
	(1.0919)	(1.1932)	(15.3375)	(1.0689)	(3.3601)	(10.0040)
R^2	0.9863	0.9942	0.9970	0.9860	0.9959	0.9975

Notes: The dependent variable is 20-year average annual GDP growth, the same as per capita growth. Robust *t*-statistics are in parentheses.

inequality. Finally, we include both inequality and financial deepening on the right hand side. Specifically, with error term ν_m for the m-th simulation or country, we run various versions of cross-country regressions:

$$Growth_m = \alpha_0 + \alpha_1 InitialFinancialDepth_m + \alpha_2 InitialGini_m + \alpha_3 Log(InitialGDP_m) + \nu_m. \tag{51}$$

Table 2 reports on these long-run, 20 year growth regressions. As shown in the first column, the higher is the "initial" 1985 level of financial deepening, the higher is the subsequent 20 year average growth rate, though the significance level is marginal. The positive sign is of course consistent with King and Levine (1993). Also, column (2) reports that the lower is the "initial" level of inequality, the higher is the subsequent 20 year average growth rate. The significance level is marginal, but the sign is consistent with Forbes (2000). However, when we include both financial depth and inequality as right-hand side variables, in column (3), the negative sign on inequality is reinforced while the sign on financial depth is reversed—both are now quite significant.[47]

[47] To the best of our knowledge, there is no empirical study that includes these two variables at the

Table 3. Spurious Growth Regression Results: Medium-Term Effects of Financial Depth and Inequality onto Growth

Estimation Method	1981–85 as initial period (1)	(2)	(3)	1976–80 as initial period (4)	(5)	(6)
Financial Depth	2.7300		1.9684	0.0599		0.6710
	(0.7645)		(0.5412)	(0.0230)		(0.2501)
Inequality		−11.2361	−10.3229		4.9854	5.4020
		(−1.2557)	(−1.1328)		(0.8192)	(0.8610)
GDP	−0.0342	1.4357	1.1718	−0.2008	−0.7181	−0.8326
	(−0.0557)	(1.2823)	(0.9574)	(−0.4596)	(−0.9936)	(−0.9728)
R^2	0.9530	0.9530	0.9531	0.9601	0.9602	0.9602

Notes: The dependent variable is 5-year average annual GDP growth, same as per capita growth. The independent variables are lagged variables. Robust *t*-statistics are in parentheses.

We repeat these regressions with 1980 as the initial year. The results on both financial depth and inequality are weakened, except for the financial depth in column (6). The substantial changes of sign, size, and significance of coefficients are consistent with the underlying transitional growth model, that is, there is no stable relationship among these variables.

Most of what shows up in the regression results appear to be determined by initial conditions and the history of shocks,[48] rather than the structure of the model. There is an extreme, if easy, way to make this point. Suppose we had taken 1976 as an "initial" period, then all 1000 simulations share the same wealth distribution in 1976, and hence the same levels of inequality and financial depth. In other words, the true initial condition is the same for all 1000 simulated economies, meaning there would not exist any meaningful relationship between initial financial depth or initial inequality and subsequent growth. The fact that this is less true over time does not mitigate the point that a

same time.

[48]Given imposed parameter values, the model generates growth in part by the shift of the population into the financial sector. As a result, levels of per-capita income, which are just the weighted averages of incomes in the participant and nonparticipant groups, are highly correlated with financial depth and hence also with inequality. This multicollinearity is a reason for the observed high R^2 in the regression.

regression of growth onto financial depth and inequality is a questionable way to think about the data and possible structural models.[49]

This instability can be observed more clearly for medium-term regressions, similar to what Forbes (2000) reports. She estimated a robust positive relationship between lagged inequality and five year average growth rates over 1965 to 1995, contrary to her long-run regressions. We construct medium-term, five year average variables[50] and conduct panel estimation of the effect of lagged financial deepening and inequality on the GDP growth rate, controlling for country fixed effects, time dummies, and the lagged GDP levels in logarithms. In sum, we run the following regression on the panel, though abusing the notation above,

$$Growth_{m,t} = \alpha_m + \alpha_1 FinancialDepth_{m,t-1} + \alpha_2 Gini_{m,t-1} + \alpha_3 Log(GDP_{m,t-1}) + \alpha_t + \nu_{m,t}, \quad (52)$$

where α_m is the country fixed effect and α_t is the time dummy.[51]

We confirm in table 3 that the regression results of medium-term panel regressions are quite different from long-run regressions. While the sign on inequality is now positive in some instances, consistent with the results of Forbes (2000), the sign is negative in other instances, depending apparently on sampling and data availability, as it were, and inequality is never significant. Indeed, none of the regressors are significant.[52]

We conclude that regressions are not an effective way to examine the data from economies in transition.

---

[49]This does not exclude the possibility of causal link from financial deepening to growth, even within our model. For example, if we generated data by varying the cost of financial access among countries (e.g., different degrees of financial repression in our model), we would see a causal relationship between financial depth/inequality and growth using the cost as an instrument in a regression. However, without the information on the cost, it would be difficult for regression studies to distinguish real effects of the institutional arrangements from spurious effects of endogenous transitional dynamics reported here.

[50]Variables are five-year average of 1976–80, 1981–85, 1986–90, 1991–95, 1996–00, and 2001–05.

[51]Note that these regressions (51) and (52) have other technical problems even if stable relationships are assumed (see Forbes (2000)). However, Forbes (2000) reports that results from simple regressions are similar to those from better estimation techniques. Given our model, even the best estimation technique would not bring meaningful results, and thus we run only these simple regressions.

[52]The high R^2 in Table 3 comes from the country-specific fixed effects, which capture the history of past shocks before the (artificial) initial period. The history of shocks up to period t determines the wealth distribution at $t-1$ and t as well as the evolution of financial participation status up to period t. As explained in section III.C, these are the drivers of the GDP growth, the financial participation rate, and Theil index.

VII. Growth, Inequality, and Financial Deepening: Examining the Model's Predictions

In this section, we use simulations to examine the model in three ways: the expected path, the best-fit simulation, and a confidence region test of the model against the actual Thai data. Our procedure is similar to what is done in the real business cycle literature, but the task here is more challenging. Because of the history dependence generated in a transitional growth path, *we need to treat the entire historical series as one sample*.

A. Expected Path

We start by examining the model based on the expected path it generates for three aggregate variables, that is, economic growth (the GDP growth rate), financial deepening (the participation rate), and inequality (the Theil index), from 1976 to 1996. Because the model does not make a sharp distinction between wealth and income, we compare the growth of wealth k in the model to the Thai GDP growth.

To generate the expected path from the model, we first construct the Markov transition matrix $\Psi(k'; k)$, defined in equation (41) for nonparticipants, and similarly for participants. We continue to assume as in the model that there is a continuum of agents so that the idiosyncratic shocks sum to zero—more generally the fraction of households with a particular set of idiosyncratic shocks is exactly the probability of that set of shocks for each household. We approximate, however, smooth wealth distributions of the continuum agents with histograms on discrete wealth grids.[53] Multiplying the approximated initial wealth distribution by the transition matrices produces the wealth distribution in the second period. The third period wealth distribution is computed in the same way as the second period wealth distribution using the same transition matrices.

Figure 4 shows the evolution of wealth distribution over time in log scale. The overall distribution becomes more dispersed over time. The lower bound is only slightly moving towards the left, due in part to low draws of idiosyncratic shocks for some households. The upper portion is expanding rapidly to the right.[54] Three aggregate variables are computed based on these wealth distributions.

[53]In the SES data, the sample size (e.g., 10,619 households in 1976) is large but of course finite. In principle, we could populate the model economy with the same number of agents with randomly drawn idiosyncratic shocks and study the evolution of the wealth distribution. However, computational issues limits us in practice. We use a smaller number of agents (1002) for the regression studies in the previous section and a continuum of agents assigned to discrete histograms in the following sections.

[54]The shapes of simulated and true distributions do not match exactly. In particular, while the actual distribution (though not shown) moved from left to right keeping a similar shape with one peak over time, the simulated distribution has an uneven shape due to a common threshold for entry across the households. This suggests a role for heterogeneity in costs, to be addressed in section VIII below. The simulated distribution also shows a much wider support than the true distribution. Despite these

Table 4. 1976–1996 Averages of Expected Paths of Macro Variables

	Growth	Fin.Dep.	Inequality
	(percentage point)		
Actual Thai Data	5.9	14.6	8.6
Benchmark Expected Path	4.0	14.8	8.6
Alternative Expected Path	6.0	19.8	9.1

Table 4 shows 1976–1996 averages of the expected paths of the three variables. At the benchmark specification for the aggregate shock θ, the model closely predicts the sample-period averages of financial deepening and inequality. However, the model predicts a low GDP growth rate relative to the actual data. With a higher mean for the aggregate shocks (the alternative case), the model replicates the sample-period average of the GDP growth rate, but over-predicts financial participation and inequality. This trade-off among the three variables appears in many of our experiments. In summary, the model can do well with sample-period averages of the three variables, but not simultaneously. It seems that there is something in the structure of the actual Thai economy which is not present in structure of the model. We return to this in the next section.

To diagnose the model's performance further, we look at entire paths, not only sample-period averages. The solid lines in figure 5 show the expected paths. While the GDP growth rate increases over time, these are low relative to the Thai data (the dashed line). In particular, the model misses the upturn in the data around 1986. Participation in the financial system, while missing the S-upturn in the data that began around 1986, goes through the middle of the Thai data, an excellent fit to the overall trend.[55] A Theil index of inequality in the bottom part of figure 5 (the solid line) shows a more or less steady increase over time with initially coinciding with the Thai data (the dashed line) up to mid-1980s, but the model prediction misses the downturn in Thai inequality in 1990s.[56]

---

two problems, the 1996 cumulative distribution of the model and that of the data are not dissimilar. Indeed, similarity is not rejected statistically in Jeong and Townsend (2001).

[55]It is not a "trend" in the strict sense defined in the business cycle literature, because again a stationarity assumption is needed to filter out trends and cycles. In this paper, we casually use the word "trend" as a linear line going through the middle of data points.

[56]Note that the initial 1976 levels of the participate rate and Theil inequality are constructed in such a way as to exactly match the data. In particular, as a result of the approximation, inequality does not necessarily match the data exactly, and we adjust the point values, so that the starting point in the model and the data are the same. There are two possible approximation errors. First, the initial wealth distribution of the model is taken as a finite approximation of the data, assigning each data

With the alternative specification, higher mean aggregate shocks, the model's prediction of the GDP growth rate, as suggested by table 4, is closer to the actual Thai data, a success. But, as is evident in figure 6, the model's prediction for financial deepening and inequality now overshoots, especially the former.

We check if the model's prediction of the observed transitional dynamics move with variations around the chosen benchmark parameters. In this way we see the mapping from the parameters of the model to policy functions and dynamic paths and so better understand the capabilities and limitations of the model. Figure 7 and 8 show two of various experiments, with policy functions on the left hand side (the dotted lines are the policy under the benchmark) and the expected paths on the right hand side (the actual Thai economy paths are again dashed lines). Lowering the variance of the idiosyncratic shock to $[-0.3, 0.3]$, as in figure 7, lowers the saving rate, causes the portfolio allocation to go to the obvious risky extreme, and raises k^* (the critical level of wealth to join the financial system). This accelerates participation well above observed Thai rates, though again the growth rate becomes more in line with the actual Thai data. A well-fit dynamic path evident in Figure 8 comes from assuming a higher return on the safe asset, to 8.49, so the share of the risky asset in the portfolio goes very low. This raises k^* considerably, and of even greater interest, captures the initially slow and subsequent fast rise in participation in the Thai data. Evidently, the prediction for participation can be improved. Raising the risk aversion parameter to 1.5 lowers the saving rate initially, though savings rises sharply with higher wealth, lowers the fraction invested in risky assets, and lowers k^* (figure is omitted). The dynamic paths of the model display slightly less growth and less participation but put the rising trend of inequality squarely through the Thai data path. Evidently, the prediction for inequality can be improved. A higher marginal transaction cost at 2 percent lowers both savings and growth. Setting a higher variation of aggregate shocks θ to 15 percent, as $\Theta = [1.022, 1.172]$, makes little changes in the expected path from the benchmark case (but see the next section, which shows a different result).

In summary, the benchmark calibrated set of parameters does a reasonable job in tracing out the trends in the data. It could perhaps do even better if we were free to choose parameters at will. In the neighborhood where we are searching, however, there are nontrivial trade-off among income growth, financial deepening, and inequality.

B. Best-Fit Simulation

We now focus on whether the model is capable of reproducing the actual economy well, if we are given the discretion of choosing a specific sequence of shocks. After all, the actual Thai path should differ from the expected path in the previous section because the actual path is imagined here to be just one realization of many possible histories of the model economy. The strategy is similar in spirit to episode studies, such as the Hayashi and Prescott (2002) study of the Japanese economy using a deterministic growth model, with TFP growth rates as measured in the national accounts.

---

point to the nearest of a large number of cells. Second, we use a log scaled grid for simulation, which more accurately approximates the true distribution for the low to medium wealth households, though less accurately for the high-wealth households.

Here, however, it is difficult to identify TFP shocks, the aggregate shocks θ in our model. A higher aggregate shock does not necessarily correspond to higher growth, because a higher shock may induce much more people to join the financial system at the end of the period,[57] resulting in possibly negative growth after netting out fixed cost q.

To identify the aggregate shocks, we select the simulation that best fits the actual Thai data among many simulations. Although it is possible to select aggregate shocks solely based on one variable (e.g., the GDP growth rate), we make use as well of the variables for financial deepening and inequality, as these are supposed to come from the same economy. Specifically, we generate 1000 simulations of the model economy, from 1976 to 1996, where each simulation is associated with a specific random draw of a sequence of aggregate shocks from the specified cumulative distribution over the 21 year period, integrating out only the idiosyncratic shocks using their known distribution. In the previous section, to compute the expected paths, all possible aggregate and idiosyncratic shocks are incorporated in the transition matrices (e.g., $\Psi(k'; k)$) through their cumulative distributions, but only the distribution of idiosyncratic shocks is used here to construct similar transition matrices given particular draws of aggregate shocks. In sum, a particular simulated path depends only on a particular sequence of aggregate shocks. Naturally, the averages of the simulated paths, averaging over different sequences of aggregate shocks (the dot-dashed lines in figures 5 and 6) are almost identical to the expected paths of these variables calculated in the previous section (the solid lines).[58]

A remaining question is how to define the "best fit" path or the measure of closeness to the actual Thai path. We let x_s denote a vector containing 57 variables—three variables (income growth, financial depth, and inequality) over 19 years (1978-1996)—for the s-th simulation, for $s = 1, \cdots, S = 1000$, and x_0 denote the actual Thai data. An obvious choice for measuring distance between each simulated path and the actual data would be $x_s - x_0$, but this is a vector, not a scaler. Note that x_s treats the entire observation in the sample period as one draw. The second choice would be the simple sum of squared errors, that is, $(x_s - x_0)'(x_s - x_0)$. However, if for example, the model generates a particular variable with high variations, then the simple sum of squared errors would be affected heavily by its performance on that particular variable. Hence, a discount for variance would be necessary. Additional discount would also be desirable for covariation of two variables, for example, income and financial deepening. If these two variables are highly correlated, then the simple sum of squared errors would be almost doubly counting the same common error in those two variables.

Hence, we measure closeness of each of the 1000 simulations (each with different aggregate shocks) to the actual Thai path using a covariance-normalized mean squared error criteria, namely,

$$(x_s - x_0)' b_0^{-1} (x_s - x_0). \tag{53}$$

---

[57]To calculate growth, we use the original timing in GJ, that is, the participation decision is made at the end of each period.

[58]Note that the expected wealth distribution (figure 4) includes dispersion caused by possible aggregate shocks. This creates a small disparity between the expected Theil indices and the average of the Theil indices of the Monte Carlo simulations.

Here covariance matrix b_0 is calculated by the sample analogue, $(1/S)\sum_{s=1}^{S}(x_s - x_0)'(x_s - x_0)$. The first two periods, 1976 and 1977, are discarded because there is little difference among simulations and the sample covariance matrix is almost singular.

The solid line in figure 9 shows the best-fit simulations from among the 1000 possibilities for the benchmark economy. The best-fit simulations show more realistic variations in the growth rate, but this variation is relatively small, not large enough to replicate the big upturn in income around 1986. At the same time, the best-fit model path shows high frequency variation that is seemingly not in the data. Strikingly, the best-fit participation rate virtually replicates the expected path and thus show no high frequency fluctuation, though still missing the S-upturn. The Theil index of inequality shows some high frequency wiggles but still misses the Kuznets-curve downturn. The alternative case with higher mean aggregate shocks shows a very similar picture (and we omit showing the plots). In short, the model can do reasonably well in delivering the trends we observe in the Thai data, since as we noted the expected path is close to the Thai trends, but not so well in delivering dynamic changes from mid-1980's.

We thus simulate another economy with a higher variation (15 percent) in aggregate shocks, that is, $\Theta = [1.022, 1.172]$ retaining the other benchmark parameter values. This makes for larger variation in the aggregate variables, as figure 10 shows, and succeeds somewhat in replicating the big upturn of the GDP growth rate around mid-1980s. There is, however, little change in the participation rate, while inequality rises faster.

C. Confidence Region

We now impose a yet more demanding standard. Namely, we test whether the actual Thai economy path itself could be a likely realization of the model at *given calibrated parameter values*. To do so, we use another mean squared error criterion, a covariance-normalized deviation of each simulated path from the the expected path a, that is,

$$\psi_s \equiv (x_s - a)'b^{-1}(x_s - a), \tag{54}$$

where b denotes the covariance matrix of the x_s around the mean a,[59] estimated from the sample as $(1/S)\sum_{s=1}^{S}(x_s - a)'(x_s - a)$.

We construct a distribution for this metric using 1000 simulation and judge if the actual Thai path x_0 lies within a reasonable range of model predictions at the calibrated parameter values. We treat the actual Thai path x_0 as one sample simulation and compute ψ_0 according to (54). We then place that test statistic among the distribution of ψ_s. We will accept the null hypothesis that the Thai path was

[59] Assuming that the model at given parameter values is the true data generating mechanism, this covariance matrix around the expected path a is the right one to use to normalize the distance statistic. In the previous section, we focused on measuring distance of a simulated path from the actual path and chose the covariance matrix around the actual path x_0 accordingly, without considering a statistical test that requires a null hypothesis.

generated under the model with its given benchmark parameter values if ψ_0 lies within the 2.5 percentile to 97.5 percentile of the density (i.e., 95 percent confidence region).

The statistic in (54) is similar to a distance statistic to be minimized in GMM estimation. There are two differences, however. First, our statistic does not necessarily follow a χ^2 distribution even with large samples. Still, as figure 11 shows, the simulated cumulative distribution of the statistic (the dashed line) is almost identical to $\chi^2(57)$ cumulative distribution (the solid line).[60] Second, while a distance statistic used in GMM estimation has the property that sample moments approach their true moments with a larger sample size, here, there is virtually zero probability that a sample path will be close to the expected path. That is, the probability density of $\chi^2(57)$ (figure 12) shows the mode of the density is 55, implying 55 is the value with the highest likelihood among possible ψ_s.[61] The expected path would by construction deliver a mean squared error of zero, but this is in the extreme left tail of the density.

The value of ψ_0 confirm that the model is not replicating all aspects of the dynamics of change of the Thai economy. Again, $\chi^2(57)$ statistic at the 2.5 percent to 97.5 percent level ranges from 38 to 80. In contrast, the Thai economy path produces a ψ_0 statistics in the far extreme right tale of the distribution. Thus, the null hypothesis that the Thai economy path could have been generated by the model at the calibrated parameter values is rejected.

We can ask nevertheless what aspects of the model fit the data best and where the fit is the worst, by decomposing ψ_0 into its elements, that is, creating covariance-normalized mean squared errors for each variable and for each combination of two variables. The model fits inequality the best, then growth, and financial deepening the worst.[62] The cross-variable covariance between growth and financial deepening is helping to reduce distances from the mean, but the opposite is true for growth and inequality.

Rejection by this strict criterion does not stem solely from the structure of the model, but also from several other sources potentially. For example, parameter values are not estimated. We know from our robustness experiments (section VII. A) that solutions can be improved on certain dimensions, though there are trade-offs among the variables.[63]

It is also possible that Thailand benefited from a rare sequence of aggregate shocks: the Thai growth episode was praised as a “miracle” in a World Bank report (1993). If this view is right, that is, the

---

[60]Tails of the simulated distribution seem thinner than the $\chi^2(57)$, but this is not always the case in our several experiments (figures are omitted). This near normality may be created by the cumulative aggregate shocks over periods. Recall that shocks themselves are drawn from a uniform, not normal, distribution.

[61]The mean is 57 and the median is 56.

[62]We suspect that the increase in wage income is associated with the decline of inequality in the Thai data. Thus asking this model to replicate the inequality path may be too much. Still, the model matches inequality with the data better than the other variables, according to our criterion.

[63]Currently we are exploring whether the techniques of this paper would allow full estimation. However, there are constraints on computation which need to be overcome.

actual aggregate shocks were outliers, then the model simulated path under such specific shock should also be an outlier. Although we cannot know the actual aggregate shocks exactly, we were able in the previous section to pick the realization that fits best to the actual Thai path. By putting the best fit simulated path into the metric (54), we can compute the relative occurrence of the best fit path among all the simulations. The best-fit simulated path at the calibrated parameter values gives 33 as the ψ_0 statistics. This test statistic is located on the left tail, 0.5 percentile, of the $\chi^2(57)$ distribution, implying the generated path is too close to the expected, smoothed path.[64] However, the test statistic for the actual Thai data is located in the right tale (far from the expected path), not in the left tail (close to the expected path). Hence, some explanation other than the realization of the extraordinary sequence of shocks is called for. Note that introducing positive temporary autocorrelation, as is with business cycle models, could help to smooth high frequency wiggles. But it would do so by generating slower movements of the variables and thus it would make the model less able to generate large, dynamic changes observed in the three macro variables.

A plausible explanation is that the anomalies are caused by a change in the financial sector policy. In fact, Abiad and others (2004) report that financial liberalization took place in Thailand from mid-1980s on and, more importantly, that under their measure of efficiency, capital allocation improved in the mid-1980s. In short, there seems to have been a policy change. While the model assumes no distortion or intervention, the path of Thailand may have been influenced in fact by this policy shift. In that case, the difference between the actual and simulated data represents the impact of changes in the government policies for the financial system.

VIII. MICRO HETEROGENEITY

It is natural to ask whether the model imposes unrealistic uniformity on the households. Indeed, the micro movements that underlie Kuznets's assertion on the aggregate statistics have been made clear in the contributions of Mookherjee and Shorrocks (1982). As with income growth, the Theil index of inequality can also be decomposed into within- and across-group changes. It consists of changes in inequality within groups, population shifts between low- and high-income groups, and changes in income differentials across these groups. These decompositions have established for several countries[65] and, for Thailand, Jeong (2000) reveals that occupation and education along with participation in the financial sector are the key driving variables. These explain[66] 72 percent of the

[64] This confirms our findings in the previous section that the best-fit path is close to the expected path and thus replicates overall trend of the Thai data.

[65] For example, in periods of nontrivial growth in Mexico and Brazil, increasing inequality (and, eventually, decreasing inequality in Brazil) are associated with changing returns to education and changing regional or urban-rural income differentials (and, in Brazil, inflation). See Bouillon, Legovini, and Lustig (1999) and Ferreira and Litchfield (1999), respectively. In Taiwan and Chile, growth coupled with an apparently stable income distribution appears to be the result of offsetting structural forces. See Bourguignon, Fournier and Gurgand (2001) and Bravo, Contreras, and Urzua (1999), respectively.

[66] Jeong (2000) uses the mean log deviation metric, a similar metric to the Theil index as both are general entropy class metric (see Cowel, 1995).

change in inequality from 1976 to 1996 (the residual is change in inequality within categories). Financial participation alone can account for 41 percent, implying it is the largest but not the unique factor.

With heterogeneity in mind, we have conducted a preliminary sensitivity analysis, specifically allowing variation in the entry cost q over different education and geographic groups, utilizing additional information in SES. Fortunately, the model with its linear returns (and no endogenous prices) allows us to calibrate and simulate for various key education and geographic groups, one at a time. For example, we can distinguish SES households by the completed level of education of the head (elementary and advanced secondary). We continue to fix technology and preference parameters at their benchmark values, including the model version of q, hence k^*. But for each group separately we center the initial (SES estimated) wealth distribution so that the initial participation rates of the Thai data match the predictions of the model (on the false assumption that everyone above the threshold was participating). Those initial participation rates in 1976 were 5 and 20 percent for the two chosen education groups (again, elementary and advanced secondary) and 5 and 16 percent for rural and urban households, respectively. We then simulate the model economy from 1976 to 1996 one group at a time.

The model at benchmark values tends to over predict subsequent participation rates of the advanced secondary and urban groups. The model predicts that access would have been higher for them over time relative to what we observe in the data (at the benchmark parameters, which, again, match reasonably well with the overall rates on average.) Ironically, one begins to wonder if there are barriers to participation, not for the poorly educated, village residents but rather for their highly educated urban counterparts. Of course we could introduce heterogeneity in preference and technology parameters across these groups, but the direction of needed change might seem surprising a priori. From our earlier robustness checks, we know that the model would require higher risk aversion for those educated/urban groups, higher variance of idiosyncratic shocks, or lower mean returns.

We also conduct another, closely related, experiment. We fix the exchange rate between the model units and Thai baht using only one group, and thus compare the transactions costs q across the various groups in the common currency units. Even though the educated have higher access, for example, they have an even higher, right-shifted distribution of wealth, so the threshold wealth k^* for them is relatively high, making their transaction costs q relatively high also. Thus we find that participation costs are higher (not lower) for the educated and urban groups. Again, to overturn this, we would need to raise (not lower) risk aversion, raise the variance of the idiosyncratic shocks, or lower the mean return for them.

In summary, there seems to emerge additional anomalies with introduction of heterogeneity. Again, one plausible explanation may lie in financial sector policies. There may have been substantial policy distortions facilitating financial access in rural areas (e.g., expansion of BAAC) while limiting access to the middle class.

IX. CONCLUSION

We follow the research agenda laid out in Lucas's Presidential Address to the American Economic Association in 2003 (Lucas, 2003):
"For us, today, value theory refers to models of dynamic economies subject to unpredictable shocks, populated by agents who are good at processing information and making choices over time. $\cdots$ [This involves] $\cdots$ formulating explicit models, computing solutions, comparing their behavior quantitatively to observed time series and other data sets. As a result, we are able to form a much sharper quantitative view $\cdots$."

In this paper, we take this agenda to the study of economies in transition, a departure from typical calibration exercises on steady-state dynamics. That is, we have proposed a quantitative research methodology consistent with the widely held view that financial deepening and changing inequality, along with economic development, are transitional phenomena. We also show that, consistent with this view, simple regression studies would not be able to capture the true linkages among growth, financial deepening, and inequality.

The model we use is a relatively simple prototype which emphasizes a fixed cost of financial participation, but this generates complicated nonlinear nonstationary dynamics. We overcome apparent nonconvexities. In particular, we point out the possible usage of a risky asset or occupation as a natural lottery which can convexify the frontier of the lifetime utility (the value function). We then establish, for each wealth level, there are uniquely determined optimal decisions, that is, savings, portfolio choices, and participation. These analytical results enable us to study the model further with numerical methods.

We apply the model to Thai data, calibrate, and make predictions. We look at the expected paths generated by the model. They are broadly consistent with the 1976–1996 averages in the data, and with the time trends of the data, especially for increasing inequality and financial deepening. By changing parameters as in our robustness checks, we can match GDP growth as well, but then actual financial depth and inequality are low compared to the model prediction. We conclude that the model is a useful starting point for studying these phenomena. Apparently we need either an additional factor promoting growth, ceteris paribus, under the benchmark parameter values, or something impeding financial deepening at higher growth as under the alternative parameters.

To examine further the model's performance along the transition, we look at the best-fit simulation, based on a covariance-normalized squared error metric of closeness. This is because the actual path of the Thai economy is imagined here to be just one realization of many possible histories of the model economy. The best-fit simulation shows some resemblance to the actual Thai data, especially to GDP growth rate. But the best-fit path at the calibrated parameters does not follow the dynamic S-curve in financial sector participation and the eventual decline of inequality. These features are statistically unlikely to be delivered by the model at the benchmark parameter values, and the actual Thai path is formally classified as outside the confidence region, constructed by a set of Monte Carlo simulations.

In sum, the model succeeds in producing the trends and some of the movements in the GDP growth, the financial access, and the Theil index evident in the Thai data from 1976 to 1996, but the model misses the specific dynamic changes that happened in the mid-1980's. We did explore several possible explanations: a rare sequence of shocks, autocorrelation in shocks, and heterogeneity in financial participation. However, these do not seem to resolve the anomalies. We conjecture there was a substantial financial sector liberalization in mid-1980s that led both to an S-curve of financial deepening and a sharp upturn in GDP growth.

The relationship among economic growth, financial deepening, and inequality are complex as they come from non-linear relationships and from an economy in transition. Simple least squares regressions do not provide the diagnostics to help us decode the meaning of the data. In this paper we have viewed both micro and macro data through the lens of a structural model. We have documented, with several diverse criteria, where the models fits well and where it falls short. Anomalies in the data relative to an explicit model have directed our thinking and can be used to guide future research. This is the agenda we commend to the reader.

Figure 2. Value Functions

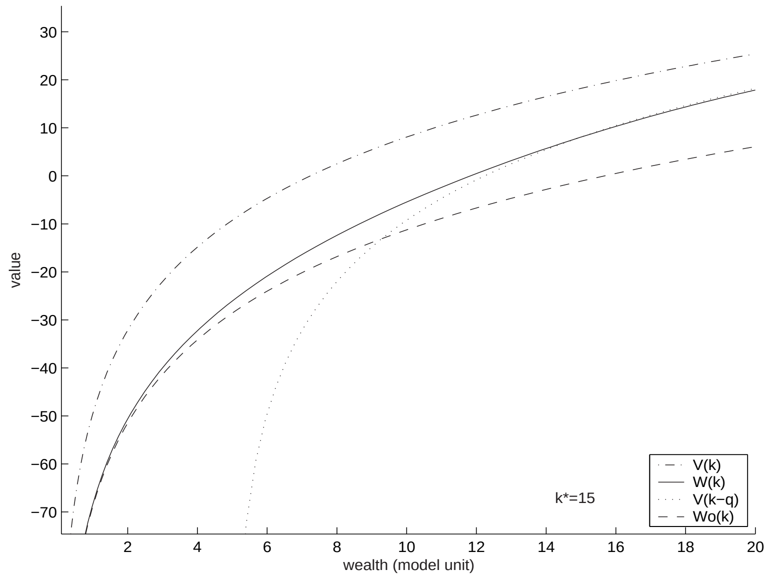


Figure 3. Policy Functions

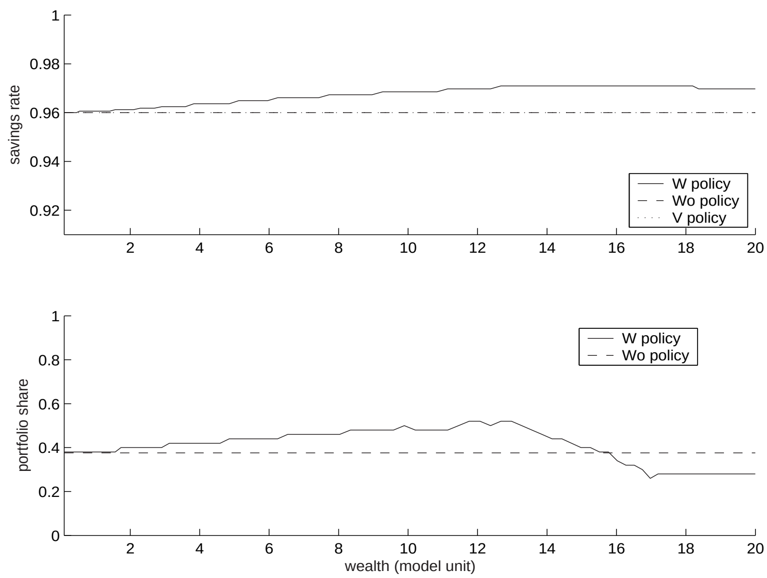


Figure 4. Wealth Evolution

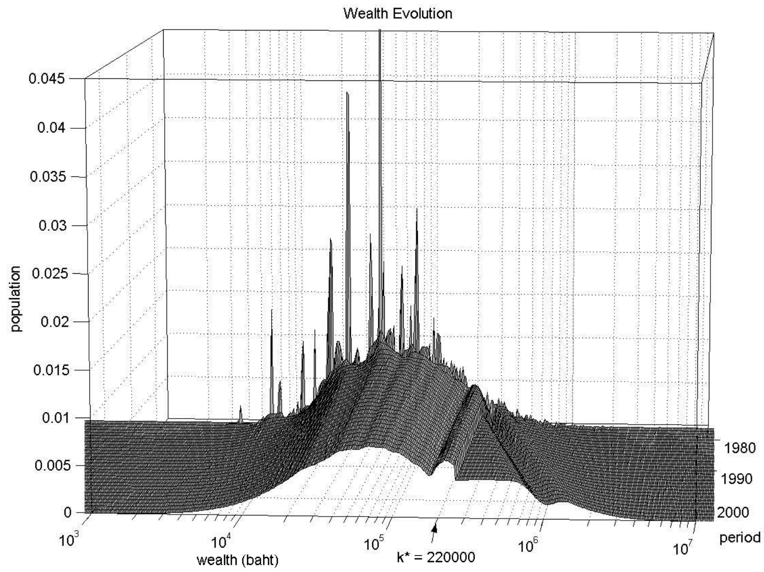


Figure 5. Benchmark Case

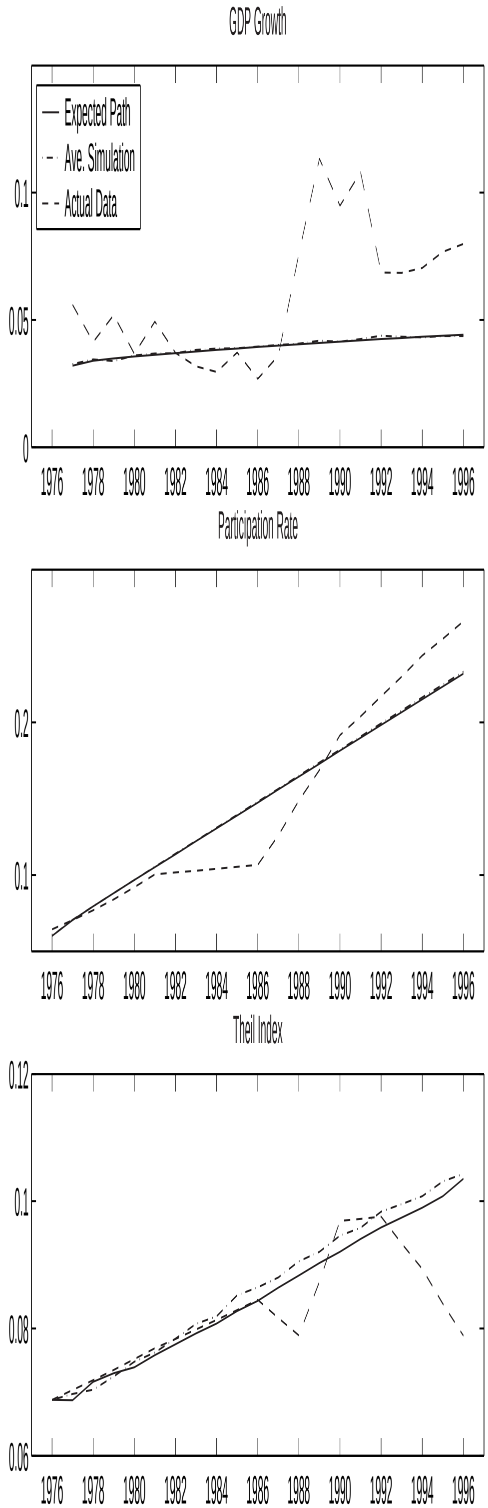


Figure 6. Alternative Case

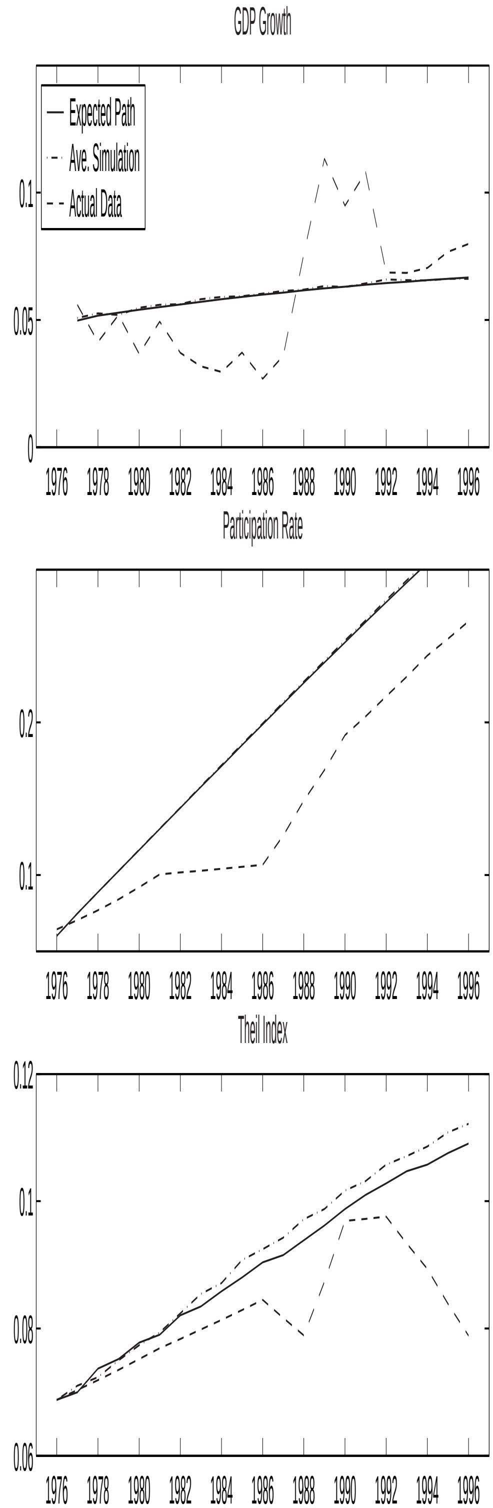


Figure 7. Lower Variance of Idiosyncratic Shock

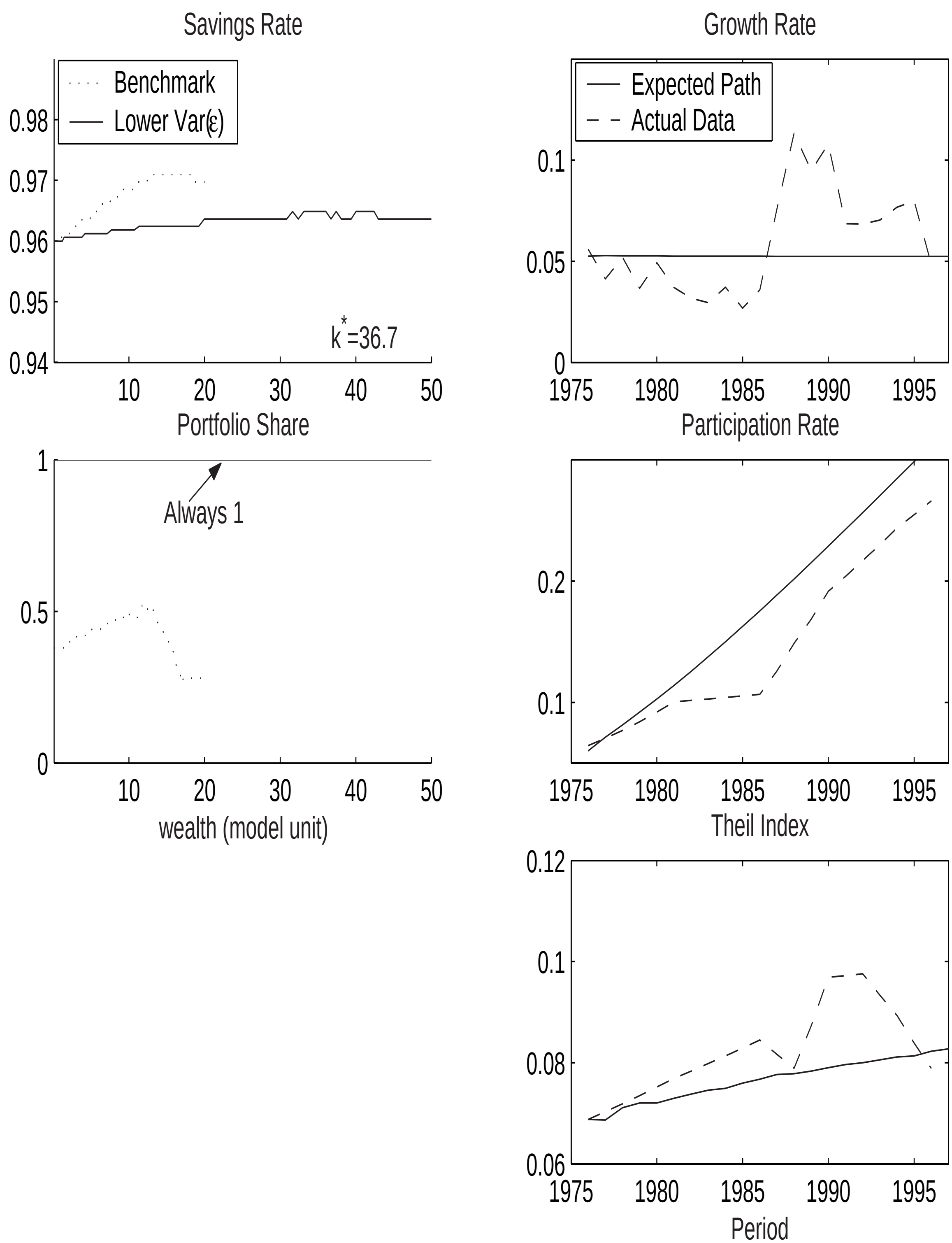


Figure 8. Higher Safe Return

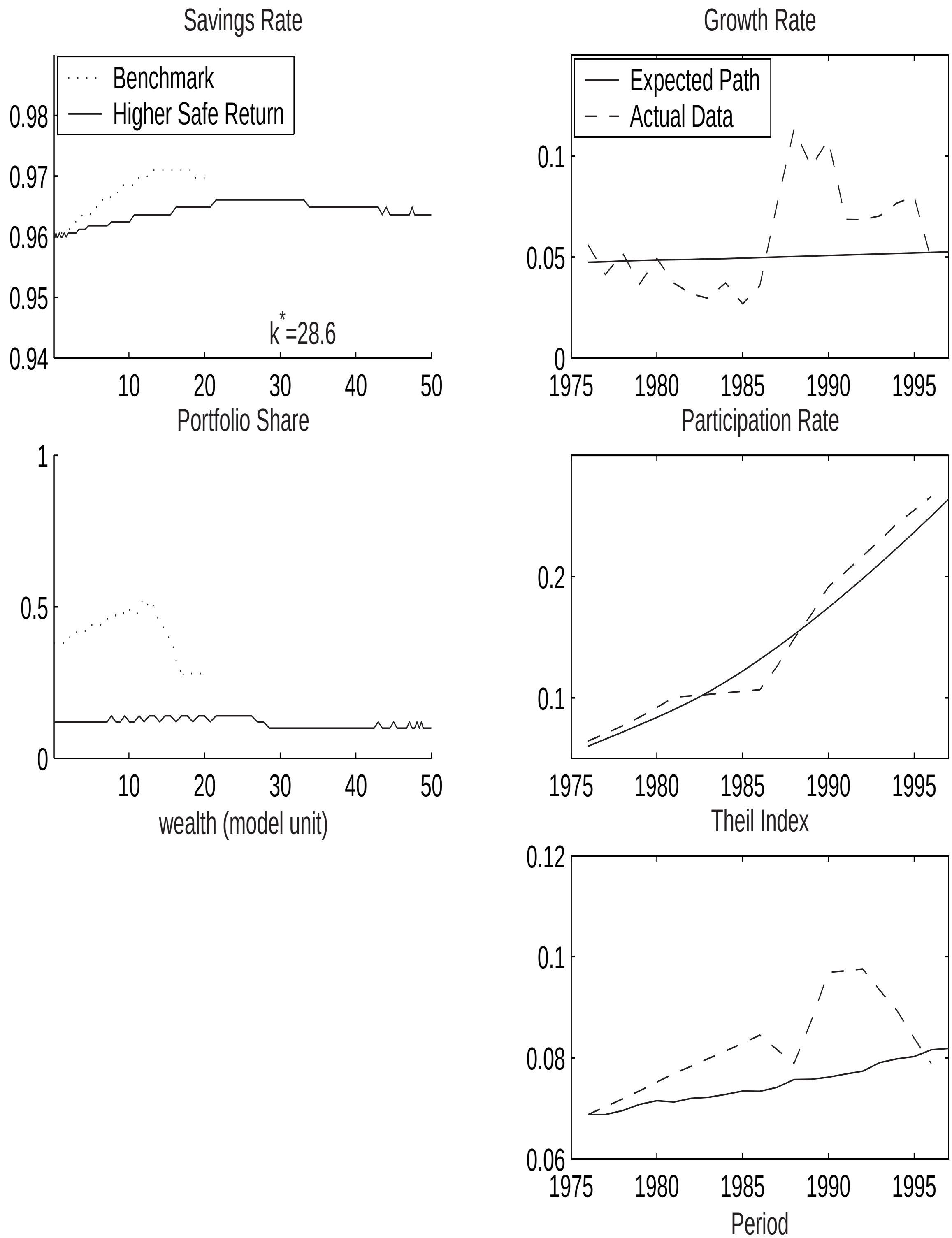


Figure 9. Benchmark, best-fit

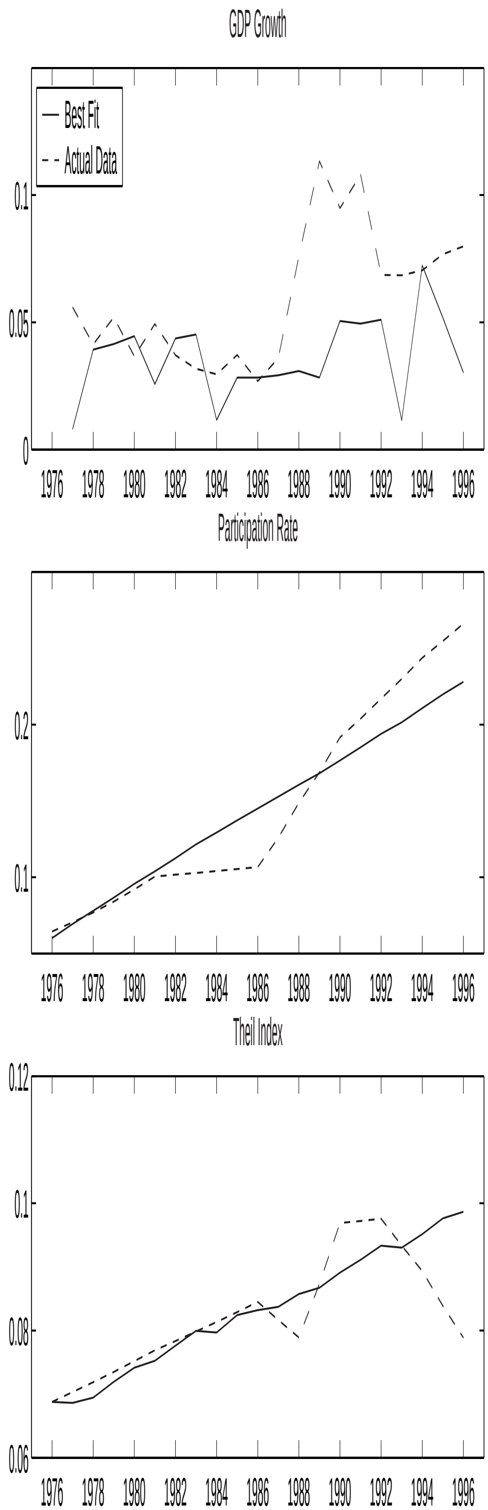


Figure 10. Higher θ Variance, best-fit

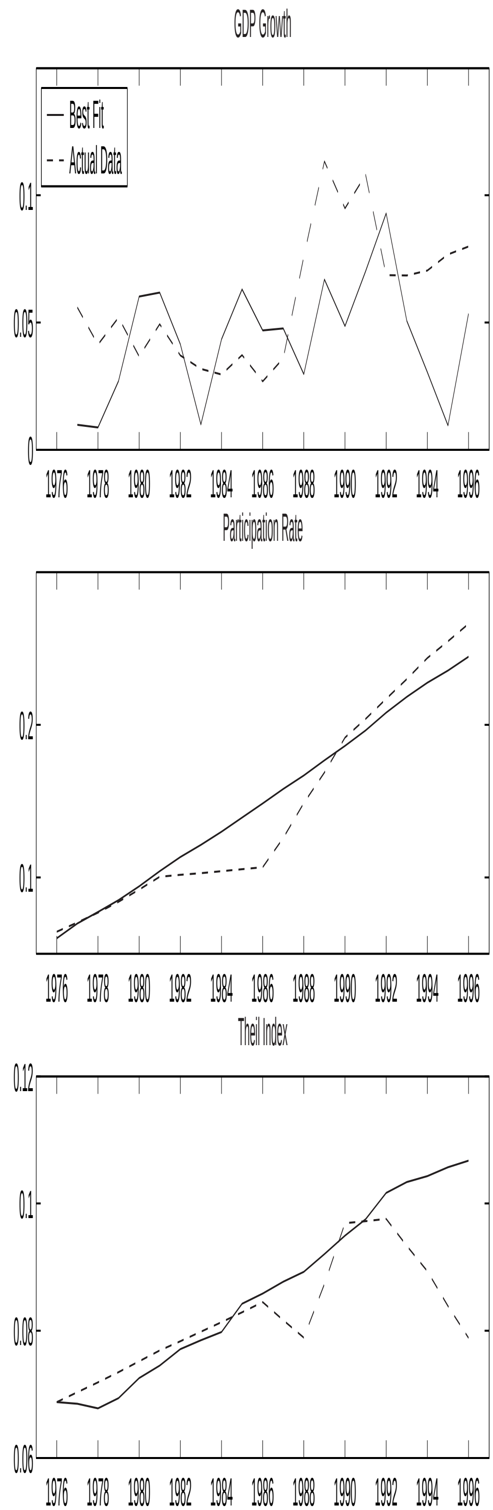


Figure 11. Cumulative Distribution of ψ_s Statistics

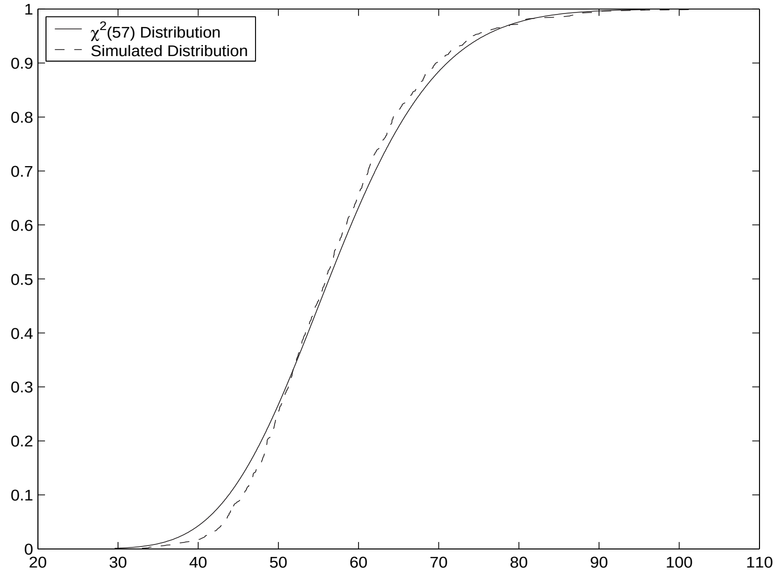
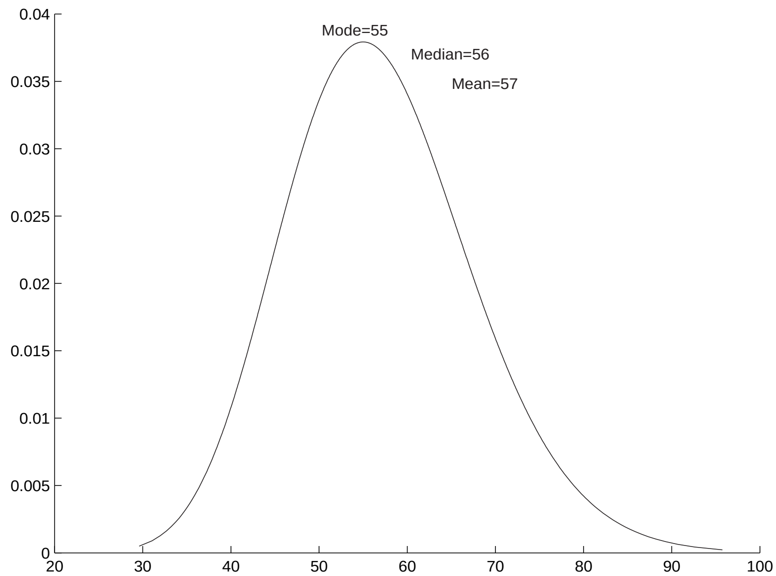


Figure 12. Probability Density of $\chi^2(57)$ Statistics



APPENDIX I. PROOFS

A. Proof of Lemma 1

Given the three properties of $Z(k)$ (i.e., single-valuedness, upper semicontinuity, and monotonicity) there are only three cases (or their combinations) that would not make $Z(k)$ a globally concave function. The first case is when $Z(k)$ jumps up at some points and is right-continuous. In this case, we can satisfy condition (18), for any sufficiently small $\xi > 0$, by taking $\tilde{k}$ as a point just before (less than ξ) the jumping point so that $\tilde{k} + \xi$ is a little larger than the jumping point. The second case is when $Z(k)$ is a continuous function on some domain but $Z(k)$ has non-convex subgraph in the neighborhood of a point, that is, $Z(k)$ is a strictly convex function in the domain of $(\tilde{k} - \xi, \tilde{k} + \xi)$. In this case, the conditions (18) and (19) are straightforward from the definition of a strictly convex function.

The final case is when $Z(k)$ is a continuous function on some domain but there is no domain $(\tilde{k} - \xi, \tilde{k} + \xi)$ on which $Z(k)$ is a strictly convex function and yet there is a domain $(\tilde{k} - \xi, \tilde{k} + \xi)$ on which $Z(k)$ has non-convex subgraph. This is possible only when $Z(k)$ is a concave function both in the domain $(\tilde{k} - \xi, \tilde{k}]$ and in $[\tilde{k}, \tilde{k} + \xi)$, while it has a kink at $\tilde{k}$ (as in figure 1). By taking any sufficiently small $\nu > 0$, we can define the left- and right-derivatives[67] at $\tilde{k}$ as

$$Z^{-}(\tilde{k}) \equiv \frac{Z(\tilde{k}) - Z(\tilde{k} - \nu)}{\nu}, \tag{A1}$$

and

$$Z^{+}(\tilde{k}) \equiv \frac{Z(\tilde{k} + \nu) - Z(\tilde{k})}{\nu}, \tag{A2}$$

respectively. To have a non-convex subgraph in $(\tilde{k} - \nu, \tilde{k} + \nu)$, the left-derivative must be smaller than the right derivative at $\tilde{k}$ (as in figure 1), implying that

$$\frac{Z(\tilde{k} + \nu) + Z(\tilde{k} - \nu)}{2\nu} > \frac{Z(\tilde{k})}{\nu}. \tag{A3}$$

By changing the notation from ν to ξ, this is equivalent to condition (18). As long as ν is appropriate choice to calculate left- and right-derivatives, any smaller $\underline{\nu} < \nu$ should also be a choice satisfying condition (A3) and, in turn, guaranteeing the condition (19).

B. Proof of Lemma 2

We can define Borel algebra on $[\underline{\eta}, \overline{\eta}]$, and for any small $\triangle > 0$, we can think of a measurable set in the neighborhood of the lowest possible shock $[\underline{\eta}, \underline{\eta} + \triangle)$ and its measure $\underline{\lambda}_{\triangle}$. Similarly, we can define a measure $\overline{\lambda}_{\triangle}$ on a measurable set in the neighborhood of the highest possible shock

[67]Since $Z(k)$ is a proper concave function in the domain $(\tilde{k} - \xi, \tilde{k}]$ and in $[\tilde{k}, \tilde{k} + \xi)$, left- and right-derivatives at $\tilde{k}$ exist (Rockafellar, 1970, pp.214).

$(\overline{\eta} - \triangle, \overline{\eta}]$. By symmetry of the probability distributions (assumption 5), these two measures should have the same mass, that is, $\underline{\lambda}_\triangle = \overline{\lambda}_\triangle = \lambda_\triangle$.

Because ξ in lemma 1 can be chosen arbitrarily small, we can always choose a portfolio share for the risky asset $\tilde{\phi}$ and an associated $\tilde{\mu}$ that equates the next period capital level with a close-to-best realization to $\tilde{k} + \xi$. That is, for any small $\triangle > 0$.

$$\left(\tilde{\phi}\left(\overline{\eta} - \frac{\triangle}{2}\right) + (1-\tilde{\phi})\delta\right)\tilde{\mu}\tilde{k} = \tilde{k} + \xi. \tag{A4}$$

By symmetry of distribution of η (assumption 5), this specific policy $(\tilde{\mu}, \tilde{\eta})$ equates the next period capital level with a close-to-worst realization to $\tilde{k} - \xi$, that is,

$$\left(\tilde{\phi}\left(\underline{\eta} + \frac{\triangle}{2}\right) + (1-\tilde{\phi})\delta\right)\tilde{\mu}\tilde{k} = \tilde{k} - \xi. \tag{A5}$$

As we discussed in the proof of lemma 1, there would be three cases when $Z(k)$ is not globally concave. The second and the third cases are when $Z(k)$ is a continuous function, and the first case is when $Z(k)$ is a discontinuous, upper semicountinuous function. In any cases, note that $Z(k)$ is Rieman integrable[68] when given compact domain, because it is real valued, single valued, monotonically increasing, upper semicontinuous, and Borel measurable.

Here, we first look at the case of a continuous function. We focus on neighborhood of $\tilde{k}$, where $Z(k)$ is not concave. Since Rieman integrability of $h(\eta)$ is guaranteed by assumption 5 (full support with positive mass everywhere), we divide its support into many grids, $j = 1, \cdots, J$ with width $\triangle$ and approximate its probability density function $h(\eta)$ by a step function having $\triangle$ as width and $h(\eta_j + \triangle/2)$ as hight, where $\eta_j \equiv \underline{\eta} + (j-1)\triangle$. Moreover we define the probability mass $D_j \equiv \triangle * h(\eta_j + \triangle/2)$. Using this notation, we can approximate a portion of integration near the worst realization by evaluating at the mid-point as follows:

$$\int_{\underline{\eta}}^{\underline{\eta}+\triangle} Z(\phi\eta + (1-\phi)\delta)\tilde{\mu}\tilde{k}dH(\eta) \approx Z\left(\phi\left(\underline{\eta} + \frac{\triangle}{2}\right) + (1-\phi)\delta\right)\tilde{\mu}\tilde{k}D_1. \tag{A6}$$

Similarly, a portion of integration near the best realization can be approximated as

$$\int_{\overline{\eta}-\triangle}^{\overline{\eta}} Z(\phi\eta + (1-\phi)\delta)\tilde{\mu}\tilde{k}dH(\eta) \approx Z\left(\phi\left(\overline{\eta} - \frac{\triangle}{2}\right) + (1-\phi)\delta\right)\tilde{\mu}\tilde{k}D_J. \tag{A7}$$

Therefore, the probability-measure weighted average of the next capital levels in the most fortunate realization and the least fortunate realizations can be approximated as a sum of two steps with associated probability mass:

$$Z\left(\left(\tilde{\phi}\left(\underline{\eta} + \frac{\triangle}{2}\right) + (1-\tilde{\phi})\delta\right)\tilde{\mu}\tilde{k}\right)D_1 + Z\left(\left(\tilde{\phi}\left(\overline{\eta} - \frac{\triangle}{2}\right) + (1-\tilde{\phi})\delta\right)\tilde{\mu}\tilde{k}\right)D_J. \tag{A8}$$

[68] See the precise definition, for example, in Stokey and others (1989).

Because $h(\eta)$ is symmetric (assumption 5), the probability measures corresponding to the highest realization D_1 and the lowest realization D_J are both equal value, denoted as λ_1. Therefore, (A8) is equal to the simple average with measure $2\lambda_1$,

$$2\lambda_1\left(\frac{1}{2}Z\left(\left(\tilde{\phi}\left(\underline{\eta}+\frac{\triangle}{2}\right)+(1-\tilde{\phi})\delta\right)\tilde{\mu}\tilde{k}\right)+\frac{1}{2}Z\left(\left(\tilde{\phi}\left(\overline{\eta}-\frac{\triangle}{2}\right)+(1-\tilde{\phi})\delta\right)\tilde{\mu}\tilde{k}\right)\right) \tag{A9}$$

Using (A4) and (A5), it is equal to

$$2\lambda_1\left(\frac{1}{2}Z(\tilde{k}-\xi)+\frac{1}{2}Z(\tilde{k}+\xi)\right). \tag{A10}$$

Thus, applying condition (18), the simple average of the outermost steps is larger than the value of Z at $\tilde{k}$, that is,

$$(A8) > 2\lambda_1 Z(\tilde{k}). \tag{A11}$$

To complete the integration for other grid between $[\underline{\eta}+\triangle,\cdots,\overline{\eta}-\triangle]$, we chose symmetrical pairs of the step functions sequentially from outside and apply the same logic above, except that we use condition (19), a weaker condition. In this way, overall integration of Z with next period capital is approximated by sum of those pair-wise approximations. Because sum of each j-th pair is larger than or equal to $2\lambda_j Z(\tilde{k})$, the overall integration has higher value than $Z(\tilde{k})$, that is,

$$\int Z(\tilde{e}(\eta)\tilde{\mu}\tilde{k})dH(\eta) > Z(\tilde{k}). \tag{A12}$$

For the discontinuous, upper semicontinuous case, to satisfy (18), we pick $\tilde{k}$ as a point just before the jumping point so that $\tilde{k}+\xi-\triangle$ is equal to the jumping point. At the same time, we pick $(\xi,\triangle)$ so that $Z(k)$ on $[\tilde{k}-\xi,\tilde{k}+\xi-\triangle)$ is a continuous function. Hence, the sum of the outermost pair would deliver the same condition as in (A11). For the remaining pairs, as long as $Z(k)$ is a convex function on the subdomain smaller than the jumping point (i.e., $[\tilde{k}-\xi,\tilde{k}+\xi-\triangle)$), the proof is the same as above.

In case when $Z(k)$ is not a convex function on the subdomain smaller than the jumping point, a sum of a pair of the step functions on the subdomain $[\tilde{k}-\xi,\tilde{k}+\xi-\triangle)$ can be smaller than $Z(\tilde{k})$, because of the possible concavity. However, as we take ξ to be very small, with associated $\triangle$, and adjusting location of $\tilde{k}$, the difference between $Z(\tilde{k})$ and the sum of a pair of the step functions on the subdomain $[\tilde{k}-\xi,\tilde{k}+\xi-\triangle)$ becomes smaller, converging to zero in the limit, because $Z(k)$ is continuous and concave function on the subdomain. But, the sum of outermost pair (A8) is significantly larger than $Z(\tilde{k})$ thanks to the discrete, finite jump at $\tilde{k}+\xi-\triangle$. Since the overall integration is the sum of all pairs of step functions, (A12) is valid in this case, too.

C. Proof of Corollary 1

We would like to apply Berge's maximum theorem and its corollary.[69] Given k_{t-1} and d_{t-1}, (μ_t, ϕ_t) can be chosen from compact, convex domain $[0,1] \times [0,1]$. To prove (μ, ϕ) are single-valued continuous function of k, it is enough to show the objective function has a range $\mathbb{R}$ (excluding $\pm\infty$) and is a strictly concave function.

First, we would like to show that the range of the objective function, $u((1-\mu)k) + \beta \int Z(e(\eta)\mu k) dH(\eta)$, is $\mathbb{R}$ given some $k \in \mathbb{R}_{++}$. Note that the solution would not change by restricting $\mu \in (0,1)$ (not taking boundary values) because of Inada conditions. This gives $|u((1-\mu)k)| < \infty$. Assumptions 3 and 4 ensure $|Z(e(\eta)\mu k)| < \infty$ for any $\eta \in \Theta + \mathcal{E}$, $\mu \in [0,1]$, $\phi \in [0,1]$, and $k \in \mathbb{R}_{++}$.[70] Hence, the objection function as a whole has a range $\mathbb{R}$.

Second, we would like to show that $u((1-\mu)k) + \beta \int Z(e(\eta)\mu k) dH(\eta)$ is strictly concave in (μ, ϕ). But u is apparently strictly concave in μ, so we only need to show concavity of $\int Z(e(\eta)\mu k) dH(\eta)$. Take any $a \in (0,1)$ as well as any (μ_1, ϕ_1) and (μ_2, ϕ_2), both from $(0,1) \times [0,1]$. Consider

$$\int Z(a(\phi_1\eta + (1-\phi_1)\delta)\mu_1 k + (1-a)(\phi_2\eta + (1-\phi_2)\delta)\mu_2 k) dH(\eta). \tag{A13}$$

Since $Z(k)$ is concave by proposition 1,

$$\begin{aligned} &\geq \int \left[aZ((\phi_1\eta + (1-\phi_1)\delta)\mu_1 k) + (1-a)Z((\phi_2\eta + (1-\phi_2)\delta)\mu_2 k) \right] dH(\eta) \\ &= a \int Z((\phi_1\eta + (1-\phi_1)\delta)\mu_1 k) dH(\eta) + (1-a) \int Z((\phi_2\eta + (1-\phi_2)\delta)\mu_2 k) dH(\eta). \end{aligned} \tag{A14}$$

Hence $\int Z(e(\eta)\mu k) dH(\eta)$ is concave in (μ, ϕ).

[69] See Stokey and others (1989) pages 62–63 and also Berge (1997) pages 116–117.

[70] See Townsend and Ueda (2001) for the proofs of boundedness the maximized life-time utility and of equivalence between the maximized life-time utility and the value function.

APPENDIX II. NUMERICAL ALGORITHM

A. Outline

The main program computes the optimal policy functions, the saving rate and the portfolio share, and the value functions. With these data, the simulation program computes the population dynamics of the economy.

The main program consists of six parts:

1. Set the relevant parameters.
2. Write the functions of $V(k)$ and $W_0(k)$ in order to refer to these values in the following procedure and to take appropriate initial function of iteration.
3. Computation of $Z(k)$.
4. Save the data of value functions and policy functions together with parameter values.
5. Simulation of population dynamics on growth and inequality using the data of 4.
6. Save the data of the simulation of 5.

B. The Construction of a Compact Domain for *Z(k)*

We use $Z(k)$ in (10) instead of $W(k)$ in (11) for iteration. Iteration on $Z(k)$ has at least two advantages over iteration on $W(k)$. One is that the $Z(k)$ formulation involves simple integration, while the $W(k)$ formulation requires an evaluation of the maximum operator inside the integrals. Essentially, the decision to join the financial intermediation is written explicitly for $Z(k)$. Simple integration saves much computational time.

The other advantage is that since $Z(k)$ takes the same value as $V(k-q)$ when k is high, we can use $V(k-q)$ as the value of $Z(k)$ for k higher than some upper end point $\overline{K}$. This is an exact extrapolation, which we do not get in the $W(k)$ formulation.[71]

The analytical results (39) and (40) suggests that those who have very small wealth act approximately as if they do not expect to join the bank ever. This implies in turn that we can truncate the domain on the left at some small capital $\underline{K}$.[72] That is, $W_0(k)$ gives us fairly accurate extrapolation value for $Z(k)$ for these lower capitals. In notation, if k_{t+1} goes lower than $\underline{K}$, $Z(k_{t+1})$ will be approximated by $W_0(k_{t+1})$. In this way, we construct a compact domain $[\underline{K}, \overline{K}]$ to compute

[71] We get the upper end-point of the domain $\overline{K}$ for computation through trial and error.

[72] Apparently, a small value of the minimum of the capital grid is better. For log utility and CRRA with $\sigma > 1$, $u(0) = -\infty$ and thus we cannot include zero in the domain of value functions. We pick 0.01 as the minimum.

$Z(k)$. Still, the value function $Z(k)$ and its policy function $(\mu(k), \phi(k))$ are defined for all $k \in \mathbb{R}_+$, for the value and policies outside the domain $[\underline{K}, \overline{K}]$ are approximated by the value functions $V(k)$ and $W_0(k)$ and their policy functions, respectively.

C. Approximation and Iteration

We use the value function iteration method to obtain values and policies. Since the model uses continuous utility functions and continuous distributions of shocks, some computational difficulties arise. The computer can only handle discrete data, and approximation of the functions and integrations are necessary.

The following is the numerical procedure to obtain value functions and policy functions.

1. First, we choose the initial given and known function $Z^0(k)$ on given $[\underline{K}, \overline{K}]$. Basically any continuous function that is between W_0 and V is appropriate.[73]

2. Second, we construct an approximation to that Z^0. This is given notationally by

$$\hat{Z}^0(k; A_0) \equiv C(Z^0(k)), \tag{A15}$$

   where C denotes the approximation procedure and A_n is the parameter of that approximation at iteration number n. Here of course $n = 0$, since we have not yet done any iteration.

   We use the Chebyshev approximation method, which is more accurate than any approximation with the same number of nodes.[74] This interpolates between special grid points by utilizing the information of all the points, and the fit is almost the best possible.

   (a) We set the Chebyshev interpolation nodes in the compact state space $k \in [\underline{K}, \overline{K}]$ for evaluating the function $Z^0(k)$. Given the degree of polynomials p and the choice of number of nodes m over $[\underline{K}, \overline{K}]$, the nodes $k(l)$ is given by

$$k(l) = (x(l) + 1)(\frac{\overline{K} - \underline{K}}{2}) + \underline{K}, \tag{A16}$$

   where $x(l)$ on $[-1, 1]$ ($l = 1, \cdots, m$, $m > p + 1$) is a Chebyshev interpolation node:

$$x(l) = \cos(\frac{2l - 1}{2m}\pi). \tag{A17}$$

   (b) Evaluate Z^0 at the nodes $k = k(l)$ for $l = 1, \cdots, m$:

$$y(l) = Z^0(k(l)). \tag{A18}$$

---

[73]See the working paper, Townsend and Ueda (2001).

[74]See Theorem 6.5.4 of Judd (1998) page 214.

(c) Then compute the Chebyshev coefficient $A_0 \equiv (A_{01}, \cdots, A_{0p})$ by the least squares method:

$$A_{0i} = \frac{\sum_{l=1}^{m} y_l T_i(x_l)}{\sum_{l=1}^{m} T_i(x_l)^2}. \tag{A19}$$

where T_i is the Chebyshev polynomial defined over [-1, 1] as

$$T_i(x) = \cos(i \arccos(x)). \tag{A20}$$

(d) Finally, we get the approximation over all k:

$$\hat{Z}^0(k; A_0) = \sum_{i=0}^{p} A_{0i} T_i(2\frac{k - \underline{K}}{\overline{K} - \underline{K}} - 1). \tag{A21}$$

3. Third, we take the appropriate extrapolation. This is for the entire range of k. Specifically, define $\overline{Z}^0(k)$ for all $k \in \mathbb{R}$ as follows.

$$\begin{aligned} \overline{Z}^0(k; A_0) &= V(k - q) \quad \text{for } k > \overline{K}, \\ &= \hat{Z}^0(k; A_0) \quad \text{for } k \in [\underline{K}, \overline{K}], \\ &= W_0(k) \quad \text{for } k < \underline{K}. \end{aligned} \tag{A22}$$

4. Fourth, we calculate $W^1(k)$ at each grid point by

$$W^1(k) = \max_{\mu,\phi} u((1-\mu)k) + \beta \int_{\underline{\eta}}^{\overline{\eta}} \overline{Z}^0(k^+(k, \mu, \phi, \eta), A_0)\, dH(\eta), \tag{A23}$$

where $k^+(k, \mu, \phi, \eta) = \mu k(\phi\eta + (1 - \phi)\delta)$.

(a) We change variables of integration. Let $h(\eta)$ denote probability distribution of η (recall the cdf was defined as $H(\eta)$). Given (k, μ, ϕ), $k^+(k, \mu, \phi, \eta)$ is a function of η. Given (k, μ, ϕ) and the specific value for k^+, η is calculated from inverse function of k^+, $\eta = k^{+(-1)}(k, \mu, \phi)$. We change the variable from η to k^+ and redefine the integrand as $\tilde{Z}(k^+; A_0)$, a function of k^+ given the Chebyshev coefficient A_0,

$$\begin{aligned} &\int_{\underline{\eta}}^{\overline{\eta}} \overline{Z}^0(k^+(k, \mu, \phi, \eta), A_0) dH(\eta) \\ &= \int_{k^+(k,\mu,\phi,\underline{\eta})}^{k^+(k,\mu,\phi,\overline{\eta})} \overline{Z}^0(k^+, A_0) h(k^{+(-1)}(k, \mu, \phi)) \frac{d\eta}{dk^+} dk^+ \\ &= \int_{k^+(k,\mu,\phi,\underline{\eta})}^{k^+(k,\mu,\phi,\overline{\eta})} \tilde{Z}(k^+; A_0) dk^+. \end{aligned} \tag{A24}$$

(b) Here, we use the Gaussian quadrature to get the approximate value of integral. The Gaussian quadrature utilizes the orthogonal polynomial approximation and calculates the integration with good accuracy and little time. Using specific discretization of k^+, which is $\{k_i^+\}_{i=1}^{p_w}$, with associated weight w_i, orthogonal approximation of the integral takes the form:

$$\int_{k^+(k,\mu,\phi,\underline{\eta})}^{k^+(k,\mu,\phi,\overline{\eta})} \tilde{Z}(k^+;A_0)dk^+ = \sum_{i=1}^{p_w} w_i \tilde{Z}^0(k_i^+;A_0). \tag{A25}$$

These $(k_i^+, w_i)_{i=1}^{p_w}$ are specific to polynomial and degree of approximation, which one can get from a table in a textbook on computation.[75]

(c) Maximization over (μ, ϕ) on equation (A23) is conducted by a grid search with successive refinements and simplex method.

5. Finally, we take the value of $Z^1(k)$ for each $k = k(l)$.

$$Z^1(k) \equiv \max_{d\in\{0,1\}} \{W^1(k), V(k-q)\}. \tag{A26}$$

Then we approximate the $Z^1(k)$ as same as $Z^0(k)$; i.e., $\hat{Z}^1(k, A_1) = C(Z^1(k))$. From this we can calculate $W^2(k)$, and construct $Z^2(k)$. This makes $\hat{Z}^2(k, A_2)$.

6. Iteration goes until $Z(k)$ converges to a fixed point.

D. Numerical Approximation of Evolution of Population Density

After the optimal policies $(\mu(k), \phi(k))$ are obtained from the numerical computation, given the initial distribution of the wealth, $M_0(k_1)$, the wealth distribution at each period for nonparticipants is recursively derived by equation (42). Here, we also need to approximate analytical distribution $\Psi(k'; k)$, equation (41), given k and k', because a computer cannot handle a continuous distribution. We use a step function approximation with finite grids. Given an initial distribution of k_1 defined on a grid, we define the distribution k_2 using the nearest point in the grid as the approximation for a particular k_2.

Note that as economy grows, the wealth distribution will disperse. This creates some difficulties that, in each tail end of distribution, the mass become tiny and cannot be captured by a computer. Here we face a trade-off, that is, making the grid space finer increases accuracy for early periods when the distribution is thick everywhere, but reduces accuracy for later periods when the distribution becomes diffused.

---

[75] See Hildebrand (1987) page 392 for a table.

References

Abiad, Abdul, Nienke Oomes, and Kenichi Ueda, 2004, "The Quality Effect: Does Financial Liberalization Improve the Allocation of Capital?," *IMF Working Paper*, No. 04/112.

Acemoglu, Daron, and Fabrizio Zilibotti, 1997, "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth," *Journal of Political Economy*, Vol. 105, No. 4 (August), pp. 709–751.

Aghion, Philippe, and Patric Bolton, 1997, "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies*, Vol. 64, No. 2 (April), pp. 151–172.

Álvarez, María J., and Antonia Díaz, 2002, "Minimum Consumption and Transitional Dynamics in Wealth Distribution," (February). Manuscript, Universidad Carlos III de Madrid.

Banerjee, Abhijit V., and Esther Duflo, 2000, "A Reassessment of the Relationship Between Inequality and Growth: Comment," (December). Manuscript, MIT.

____ , and Andrew F. Newman, 1993, "Occupational Choice and the Process of Development," *Journal of Political Economy*, Vol. 101, No. 2 (April), pp. 274–298.

Bencivenga, Valerie, and Bruce D. Smith, 1991, "Financial Intermediation and Endogenous Growth," *Review of Economic Studies*, Vol. 58.

Berge, Claude, 1997, *Topological Spaces* (Mineola: Dover).

Billingsley, Patrick, 1995, *Probability and Measure* (New York: John Wiley & Sons, Inc.).

Bouillon, Cesar B., Arianna Legovini, and Nora Lustig, 1999, "Rising Inequality in Mexico: Returns to Household Characteristics and the 'Chiapas Effect'." Manuscript, Inter American Development Bank.

Bourguignon, Francis, Martin Fournier, and Marc Gurgand, 2001, "Fast Development with a Stable Income Distribution: Taiwan, 1979–94," *Review of Income and Wealth*, Vol. 47, No. 2 (June), pp. 139–163.

Bravo, David, Dante Contreras, and Sergio Urzua, 2002, "Poverty and inequality in Chile 1990–1998: Learning from Microeconomic simulations," *Universidad de Chile Working Paper*, No. 196 (October).

Cowell, Frank, 1995, *Measuring Inequality* (Englewoods Cliffs, New Jersey: Prentice-Hall).

Danforth, John P., "On the Role of Consumption and Decreasing Absolute Risk Aversion in the Theory of Job Search," 1979, in Steven A. Lippman, and John J. McCall, eds., *Studies in the Economics of Search*, (New York: North-Holland), pp. 109–131.

De Nardi, Mariacristina, 2004, "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, Vol. 71, No. 3 (July), pp. 743–768.

Deaton, Angus, 1997, *The Analysis of Household Surveys: A Macroeconometric Approach to Development Policy* (Baltimore: Johns Hopkins University Press).

Den Haan, Wouter J., and Albert Marcet, 1994, “Accuracy in Simulations,” *Review of Economic Studies*, Vol. 61, No. 1 (January), pp. 3–17.

Favara, Giovanni, 2003, “An Empirical Reassessment of the Relationship Between Finance and Growth,” *IMF Working Paper*, No. 03/123.

Ferreira, Francisco H.G., and Julie A. Litchfield, 1999, “Education or Inflation? The Roles of Structural Factors and Macroeconomic Instability in Explaining Brazilian Inequality in the 1980s,” *STICERD Discussion Paper*, (March). DARP 41.

Forbes, Kristin J., 2000, “A Reassessment of the Relationship Between Inequality and Growth,” *American Economic Review*, Vol. 90, No. 4 (September), pp. 869–887.

Goldsmith, Raymond W., 1969, *Financial Structure and Development* (New Haven: Yale University Press).

Gomes, Joao, Jeremy Greenwood, and Sergio Rebelo, 2001, “Equilibrium Unemployment,” *Journal of Monetary Economics*, Vol. 48, pp. 109–152.

Greenwood, Jeremy, and Boyan Jovanovic, 1990, “Financial Development, Growth, and the Distribution of Income,” *Journal of Political Economy*, Vol. 98, pp. 1076–1107.

Hayashi, Fumio, and Edward C. Prescott, 2002, “The 1990s in Japan: A Lost Decase,” *Review of Economic Dynamics*, Vol. 5 (January), pp. 206–235.

Hildebrand, Francis B., 1987, *Introduction To Numerical Analysis* (Mineola: Dover).

Jeong, Hyeok, 2000, “Sources of Kuzunets Dynamics in Thailand.” PhD dissertation, (Chicago: The University of Chicago).

____ , and Robert M. Townsend, 2001, “Evaluation of Models of Growht and Inequality.” Manuscript, The University of Chicago.

Judd, Kenneth L., 1998, *Numerical Methods in Economics* (Cambridge: MIT Press).

King, Robert G., and Ross Levine, 1993, “Finance and Growth: Schumpeter Might be Right,” *Quarterly Journal of Economics*, Vol. 108, No. 3 (August), pp. 717–737.

Klinhowhan, Ubonrat, 1999, “Monetary Transmission Mechanism in Thailand,” Master’s thesis, (Bangkok: Thammasat University).

Krusell, Per, and Anthony A. Smith, Jr., 1998, “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, Vol. 106, No. 5 (October), pp. 867–896.

Kuznets, Simon, 1955, “Economic Growth and Income Inequality,” *American Economic Review*, Vol. 45, No. 1 (March).

Kydland, Finn, and Edward C. Prescott, 1982, “Time to Build and Aggregate Fluctuations,” *Econometrica*, Vol. 50, No. 6 (November), pp. 1345–1370.

Lloyd-Ellis, Huw, and Dan Bernhardt, 2000, “Enterprise, Inequality and Economic Development,” *Review of Economic Studies*, Vol. 67, No. 1 (January), pp. 147–168.

Lucas, Jr., Robert E., 2003, “Macroeconomic Priorities,” *American Economic Review*, Vol. 93, No. 1 (March), pp. 1–14.

McKinnon, Ronald I., 1973, *Money and Capital in Economic Development* (Washington, DC: Brookings Institution).

Mookherjee, Dilip, and Anthony F. Shorrocks, 1982, “A Decomposition Analysis of the Trend in UK Income Inequality,” *Economic Journal*, Vol. 92, No. 368 (December), pp. 886–902.

Mortensen, Dale T., 1989, “The Persistence and Indeterminacy of Unemployment in Search Equilibrium,” *Scandinavian Journal of Economics*, Vol. 91, No. 2, pp. 347–370.

Piketty, Thomas, 1997, “The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing,” *Review of Economic Studies*, Vol. 64, No. 2 (April), pp. 173–189.

Saint-Paul, Gilles, 1992, “Technological Choice, Financial Markets and Economic Development,” *European Economic Review*, Vol. 36, No. 4 (May), pp. 763–781.

Shaw, Edward S., 1973, *Financial Deepening in Economic Development* (New York: Oxford University Press).

Stokey, Nancy L., Robert E. Lucas Jr., and Edward C. Prescott, 1989, *Recursive Methods in Economic Dynamics* (Cambridge: Harvard University Press).

Townsend, Robert M., 1978, “Intermediation with Costly Bilateral Exchange,” *Review of Economic Studies*, Vol. 45 (October), pp. 417–425.

____, 1983, “Financial Structure and Economic Activity,” *American Economic Review*, (December), pp. 895–911.

____, Anna Paulson, Sombat Sakuntasathien, Tae Jeong Lee, and Mike Binford, 1997, “Questionnaire design and data collection for NICHD grant *Risk, Insurance and the Family*, and NSF grants.” Manuscript, The University of Chicago.

____, and Kenichi Ueda, 2001, “Transitional Growth with Increasing Inequality and Financial Deepening,” *IMF Working Paper*, No. 01/108.

World Bank, 1993, *The East Asian Miracle: Economic Growth and Public Policy* (New York: Oxford University Press).