Macroeconomic Dynamics, 1–54. Printed in the United States of America.
DOI: 10.1017/S1365100507070149

GROWTH AND INEQUALITY: MODEL EVALUATION BASED ON AN ESTIMATION-CALIBRATION STRATEGY

Hyeok Jeong
University of Southern California

Robert M. Townsend
University of Chicago

This paper evaluates two well-known models of growth with inequality that have explicit micro underpinnings related to household choice. With incomplete markets or transactions costs, wealth can constrain investment in business and the choice of occupation and also constrain the timing of entry into the formal financial sector. Using the Thai Socio-Economic Survey (SES), we estimate the distribution of wealth and the key parameters that best fit cross-sectional data on household choices and wealth. We then simulate the model economies for two decades at the estimated initial wealth distribution and analyze whether the model economies at those micro-fit parameter estimates can explain the observed macro and sectoral aspects of income growth and inequality change. Both models capture important features of Thai reality. Anomalies and comparisons across the two distinct models yield specific suggestions for improved research on the micro foundations of growth and inequality.

Keywords: Growth and Inequality, Wealth-Constraints, Self-Selection, Occupational Choice, Financial Deepening

1. INTRODUCTION

Our purpose is to understand growth and inequality. We do this through an evaluation of macro models that are explicit about micro underpinnings and impediments to trade. We use the explicit structure of the macro models to make exact numerical predictions for aggregate dynamics, the dynamics of key subgroups, and end-of-sample period income distributions. We compare these predictions to those objects in the data from a given, selected country. In this sense, we take theory seriously as

We thank Natalia Ramondo, Alex Karaivanov, and David Cuberes for research assistance. We also appreciate the helpful comments from Lars Hansen, Hugo Hopenhayn, Roger Moon, and the participants of the conference in Buenos Aires honoring Rolf Mantel, European Meeting of Econometric Society in Venice, and the Money and Banking workshop and Applications of Economics workshop at the University of Chicago. Financial support from NSF (SES-0318340) and NICHD (R01 HD27638) is gratefully acknowledged. Address correspondence to: Hyeok Jeong, Department of Economics, University of Southern California, 3620 S. Vermont Ave., Kaprielian Hall Room 300, Los Angeles, CA 90089, USA; e-mail: hjeong@usc.edu.

 1365-1005/07 $18.00

in the calibration, real business cycle literature, pioneered by Kydland and Prescott (1982). However, the parameters of preferences, technology, and distribution of shocks that we use are neither arbitrarily chosen nor borrowed from other studies. Rather, we explicitly estimate the key parameters using the models' micro underpinnings, which presuppose that household and business choices are constrained by wealth. Specifically, we use repeated cross-sections of micro data from the country that we study to estimate most of these key parameters, using maximum likelihood methods on the household and business choices.

In the growth and inequality literature, tight links between theory and data are rarely pursued. Various equilibrium theories of growth and inequality have been developed without evaluating their empirical validity. Until recently, studies of the empirical relationship between growth and inequality are mostly based on cross-country regression analysis. The results of these reduced-form cross-country studies are, unfortunately, robust neither to the specification of estimation nor to the selection of data, and have provided only suggestive clues.[1] A more natural alternative for studying the dynamic relationship between growth and inequality is an analysis of the evolution of the income distribution for a given country over time, using a series of micro data. Bourguignon (2002) makes an excellent case.

Our goal is both methodological and positive. The methodological goal is to accomplish what we feel is a natural synthesis between theory and econometrics. On the positive side, we use the method to discover where the models fit well and where salient anomalies appear, to guide the construction of new and better models. More generally, we hope that our model evaluation exercise might advance the "mutual penetration of quantitative economic theory and statistical observation," as Frisch (1933) envisioned. In particular, we hope that this mutual penetration helps to narrow the gap between "theory without measurement" and "measurement without theory" in the growth-inequality literature.

Two immediate comments are in order. First, our merging of micro with macro in both theory and estimation is sufficiently challenging that we try to stick to familiar ground on other dimensions. Namely, we use two models that are reasonably well known and presumably well understood and analyze each of the models against micro and macro data as they stand, rather than leaping to an integrated model that combines in one spot the salient features of the two models. The paper does conclude with specific instructions for future model construction, incorporating lessons learned from the model evaluation.

Second, we use neither the macro aggregate dynamics of growth and inequality nor the shape of the income distribution in our estimation. These data are saved in order to compare the models' macro dynamics predictions with those in the data. We use model's implications of individual agents's choices only to fit the micro data. This two-step procedure, setting aside some of the data for "verification" or "testing," and excluding its use in "parameter selection," is a commonly established practice in empirical work in the natural sciences. See Oreskes, Shrader-Frechette, and Belitz (1994) and Hansen and Heckman (1996), for example, who provide a lively discussion about this two-step procedure as an empirical strategy.

Our empirical strategy is based on the following reasoning. First, the separation between verification and parameter selection helps us to avoid "overfitting" of the model to the data, as Granger (1999) discusses. Second, we fit to the micro data of household choices, not to the aggregate dynamics of growth and income distributions, because we consider household choices as the (micro) foundation of the objects that we want to explain, that is, macro dynamics of growth and inequality.[2]

Here we apply our method to Thailand, a country that grew rapidly for the 1976–1996 period, but with increasing and then decreasing inequality. Real GNP per capita grew at 5.7% annually. In particular, for the 1986–1996 period, the average annual growth rate at 9% even exceeded those of neighboring East Asian miracle economies. However, the already high-income Gini coefficient of Thailand at 0.436 in 1976, close to the average-income Gini coefficient of sub-Saharan African countries at 0.441, increased to 0.515 by 1996, exceeding the average income Gini coefficient in Latin American and Caribbean countries at 0.502.[3]

Furthermore, we know from Jeong (in press) and from his use of the Thai Socio-Economic Survey (SES) data that substantial parts of this growth with changing inequality are accounted for by population shifts across subgroups and associated changes in income gaps. As population shifts across sectors were postulated by Kuznets (1955) to be the driving force of the relationship between growth and income inequality, the focus on self-selection as a micro foundation of growth and inequality is natural. Two key characteristics that account for the changes are occupational shift and financial deepening. Thus, we adopt from the literature two reasonably well-known models of growth and inequality that emphasize one or the other of these self-selection components, although there are other possible channels for growth and inequality.[4] In the occupational choice model of Lloyd-Ellis and Bernhardt (2000) (hereafter LEB), households of varying talent face imperfect credit markets in financing occupational choice and the scale of enterprise. Thus households are constrained by limited wealth, although this can be alleviated over time. As the distribution of wealth evolves, so do the occupational composition of population and income differentials, generating the dynamics of growth with changing inequality. Likewise, in the financial deepening model of Greenwood and Jovanovic (1990) (hereafter GJ), households face wealth constraints in their decisions to undertake costly entry into the financial system itself. Participation in financial intermediaries provides the benefits of sharing idiosyncratic risks and advanced information on aggregate risks. As economy-wide wealth shifts to the right, more households gain access to financial intermediaries, and this in turn affects growth and inequality dynamics.

Until recently, neither model has been brought to actual data. Exceptions are Gine and Townsend (2004) who estimated the LEB model and Townsend and Ueda (2006) who calibrated the GJ model. Both papers study the aggregate implications of the models. Here we study the decomposed subgroup dynamics and the shapes of the implied income distributions as well as the aggregate dynamics of growth and inequality. We also implement structural estimation using the explicit dynamic programs of the participants and nonparticipant subgroups of the GJ model. This is different from the calibration exercise of Townsend and Ueda (2006). Furthermore, our comparative evaluation of the two models with the same data addresses in another way their empirical validity, and yields the instructions for future model construction.

Section 2 provides an overview of Thai growth and inequality, and summarizes the results of the model evaluation exercise. Section 3 describes the occupational choice model including estimation, simulation, and its fit to macro dynamics, subgroup dynamics, and end-of-sample-period income distribution. Section 4 follows the same ordering for the financial choice model. Section 5 concludes.

2. OVERVIEW

2.1. Thai Background

Thai per capita income grew at 4.5% each year on average for the two decades between 1976 and 1996, with low or slightly negative growth from 1976 to 1986, surging to 12% in late 1980s, and then declining to 6% by 1996. The LEB model economy captures the recession, surge, and decline, although the magnitudes are smaller, for example, from 1% to 7% in the surge, and a decline to 2% after. Thai income inequality (measured by Theil-L index) also grew at 2.4% each year and displays a Kuznets curve, first increasing from 0.39 in 1976 to 0.66 in 1992, and then declining to 0.58 by 1996. The model mirrors this movement in inequality quite well, although it underpredicts the level throughout, specifically increasing from 0.24 in 1976 to 0.50 in 1992, with a more dramatic and earlier decline to 0.33 by 1996. The Thai fraction of the population who are either self employed or hire workers was virtually constant around 15% for the period of 1976–1990 and then increased to 19% during 1991–1996. The model economy captures these dynamics as well, although again the levels are lower.

The fraction of wage earners who are not in the financial sector declined steadily in the Thai economy from 80% to 60%, and the LEB model economy imitates this decline, from 90% to 70%. The fractions of financial sector wage earners and entrepreneurs rise from 5% to 21% and 2% to 6%, respectively, almost exactly as in the model. There are in the data many entrepreneurs among those participating in the financial sector, about 21% on average, although in the model this is higher, 28% on average. But in the data, 15% of those in “autarky” (i.e., nonfinancial sector) are also entrepreneurs, whereas it is only 2% in the model.

Wage earners who do not participate in the financial sector earn the least and their earnings are more or less constant until late 1980s, but their earnings increase in the 1990s. The model imitates this pattern. The highest-earning group in Thailand are financial sector participants, especially those running enterprises, whereas the highest-earning group in the model economy are entrepreneurs, especially nonparticipants. There is also much co-movement in earning growth across all sectors in Thailand, whereas in the model, movements of the wage and interest rate cause divergence. In Thailand, the entrepreneurial premium, the ratio of profits relative to wages, is high at almost 2 in 1976, and this declines to 1.6 by 1996. In the model, the premium is overdone and the decline even more dramatic, from 20 to 8.

Inequality in Thailand, by 1986 and beyond, is highest among participant entrepreneurs, ranging overall from 0.45 to 0.70. In the model, that group also has the highest level of inequality, but the level is lower at 0.15 on average. The dynamics of aggregate inequality is captured well by the model via changes in across-group inequality.

In Thailand, the population shifts from wage earners to entrepreneurs, and from nonparticipants in financial sector to participants, contribute 1.6% of the total 4.5% of annual per capita income growth. This leaves 64% of growth attributed to within-subgroup growth. Occupational shifts alone, ignoring changes in financial participation status, account for one-tenth of the compositional effect on growth. The joint compositional effects on growth in the model is larger at 3.6%.

Of the total inequality growth rate at 2.4% per year in Thailand, 0.6% is a result of population shifts into higher income categories, −0.02% because of income convergence across subgroups, and 0.14% because of population shifts into higher inequality categories. The rest, 1.7%, is from the inequality changes within subgroups. Put differently, in Thailand 70% of inequality change is not accounted for by these occupation/financial sector categories. In the model, the signs are correct, but the orders of magnitude are not. Of the total predicted inequality growth at 1.7%, 6.9% is because of population shifts into higher income categories, −3.8% because of income convergence across subgroups, 0.3% because of population shifts into higher inequality categories, and 0.1% because of the inequality changes within subgroups. Essentially, the majority of the inequality change in Thailand is within categories, whereas in the model, it is across categories.

The LEB occupational choice model predicts much higher concentration at the lower left tail of the income distribution and hence more poverty than in the data. Because of this difference, Kolmogorov-Smirnov test of goodness of fit for the end-of-sample-period income distribution rejects the resemblance of the shape of income distribution between the LEB model and the Thai data.

The Thai financial sector deepened in the sense that the population fraction of the participants in the formal financial intermediaries increased from 6.4% to 26.6% for the two decades, in particular with nonlinear acceleration during the 1986–1996 period. The GJ financial deepening model gets the trends of financial deepening, income growth, and inequality increase almost exactly. However, the model misses some dynamics of the data, for example, the downturn in inequality after 1992 and the 1986–1996 surge in financial deepening.

The income levels of Thai financial sector participants are higher than nonparticipants, at an initial ratio of 2.3, rising to 3.1 in 1992, then falling back to 2.4 by 1996. In the financial deepening model, the gap is continuously increasing, and the premium is overdone, rising from 5 to 7. There is much more co-movement of growth between the financial and nonfinancial sectors in the data than in the model, despite aggregate shocks in the model. Inequality is lower among participants only during 1979–1987 in Thailand, whereas in the model inequality among participants is lower than the nonparticipants throughout, rising from 0.02 to 0.2.

In contrast to the small effects of occupational transition on growth, financial deepening contributed to Thai income growth at 1.6% each year, 36% of total annual growth. Its effects on Thai income inequality also were substantial, contributing to 0.83% of inequality growth each year, 34% of the total inequality growth (combining the effects resulting from population shifts into higher-income categories and into higher-inequality categories). The GJ model predicts substantial effects of financial deepening on growth and inequality change as in the data but too large in order of magnitudes, 3.3% for income growth, 1.3% for inequality increase. Again, the major source of the Thai inequality increase is a result of the increases within financial and nonfinancial sectors, 1.52% of increase each year (63% of the total inequality growth). This effect in the GJ model is tiny at −0.1%.

The model misses some variety among the poor, misses the very rich, and is bimodal in predicting the 1996 Thai income distribution, but overall the predicted income distribution is not statistically different from the actual Thai income using the same Kolmogorov-Smirnov test statistic.

2.2. Comparative Model Evaluation

The underlying mechanisms driving growth are different between the two models. GJ is an *Ak* type of growth model with aggregate shocks, and LEB is a typical neoclassical growth model subject to diminishing returns without aggregate shocks. The key selection characteristics are also different, occupational choice in LEB and financial participation in GJ. Nevertheless, at parameter values that best fit the assumed micro selection decisions, both models predict the aggregate dynamics of growth and inequality change reasonably well.

Of course, there are remaining differences between these models. GJ fits the inequality level better than LEB. But movements over time of average income and income inequality are better captured by LEB than GJ, without aggregate shocks, through endogenous factor price movements. Evidently, the modeling of endogenous movements of wages and interest rates is important. However, we would not underplay the importance of aggregate shocks. In fact, there are strong co-movements in growth rates across subgroups in the data (except the catchup growth periods in the early 1990s). But both models fail to capture these co-movements. Again, LEB simply lacks aggregate shocks. GJ does have aggregate shocks in the risky technology, but the effects are asymmetric between nonparticipants and participants. The financial sector can mitigate the adverse effects of aggregate shocks as they are foreseen. Nonparticipants are exposed to them, yet they diversify into the low-yield safe technology. Thus in the end, co-movements in growth rates across participant and non-participant groups are weak. Evidently, we need another common factor that affects all subgroups, or we need to make the informational advantage of the financial sector in forecasting aggregate shocks less perfect.

Both models predict population shifts across key selection groups, from low-income groups to high-income groups (entrepreneurs in LEB and financial sector participants in GJ), and the effects of these compositional changes in the population on growth and inequality change are substantial, especially for financial participation, as in the data. Thus, we find that wealth-constrained choices related to occupation and financial participation are indeed significant channels at the micro level that link growth and inequality dynamics at the macro level. This confirms the existence of an intimate relationship between growth and inequality, as Kuznets (1955) postulated. Here we identify occupational choice and financial sector participation as sources of the relationship.

Differential growth rates across subgroups is another channel linking growth and inequality at the aggregate level. In LEB, the income of the rich group, entrepreneurs, declines over time because of the built-in diminishing returns, whereas that of the poor group, workers, is increasing. Thus, positive within-group growth comes only from the poor group. Income inequality eventually decreases via this catchup effect. In contrast, in GJ, income grows faster in the rich group, participants in the financial sector, than in the poor group, nonparticipants. The income gap between two groups diverges. This is one of the main sources of increasing inequality in GJ. Likewise, catchup growth is not captured by GJ. Poor households either suffer low growth or graduate into the financial sector. In both models, overall within-group growth is low relative to the data. There seem to be some missing engines of growth within subgroups. Human capital is a candidate, as identified by Jeong (2000).

Both models predict income gaps across key selection groups, which are far too large. Thus, the orders of magnitude of the composition effects on income growth and inequality change are exaggerated relative to the data. For the same reason, the convergence effect in LEB and the divergence effect in GJ on inequality change are also exaggerated. Each model has two kinds of heterogeneity: individual wealth and random entrepreneurial talent in LEB and individual wealth and idiosyncratic shocks for the risky investment in GJ. The assumed cross–sectional variation of occupational choices or financial participation choices with wealth, when they are forced onto the structure of the model, seem to require large counterfactual income gaps.

Remarkably, the GJ model predicts well the overall shape of the income distribution at the end of the sample period. But both models do not predict well-observed income variation in the upper and lower tails of income distribution. That is, for upper tails, both models fail to predict the existence of extremely rich people. As for the lower tail, the models fail in different ways. In LEB, wage earners and subsisters are all alike, which creates a spike at the low end of income distribution and hence no income variation among poor people. In GJ, the poor group, nonparticipants in financial sector, are subject to idiosyncratic shocks. There is reasonable income variation among them, but the lower tail is strictly shifted to left, that is, there are too many poor people relative to the data. A mechanism is needed in GJ under which poor people can escape poverty even when they are outside the financial system. LEB with its increasing wages provides one such mechanism.

There is also a noticeable discrepancy in subgroup inequality levels. In the data, financial sector participant and nonparticipant groups have similar levels of inequality. In GJ, the inequality level is much lower among participants than among nonparticipants. But in LEB, the opposite is true. The low level of inequality among participants in GJ suggests that financial sector is too good at diversifying idiosyncratic shocks. Thus, we may need to make the insurance role of the financial sector less than perfect. In LEB, the level of inequality is higher for participants because of high interest rates (the return on saving amplifies wealth differences into income differences), and because of high variation in wealth and talent for participants (as low-wealth but high-talent households are not constrained in setting up business in the financial sector but need to pay back interest). This suggests that making the credit market less than perfect will improve subgroup inequality dynamics. It is interesting that two quite distinct models suggest the common necessity of less-than-perfect financial markets.

Indeed, one way to think about the model comparison exercise is to begin with LEB, which has an exogenously expanding financial sector, and then compare it to GJ with its endogenous costly entry. LEB is at loss to explain the higher entrepreneurial income of financial sector participants, but GJ can reconcile that anomaly. In effect, wealth constraints (entry cost) create rents, which shows up as income differentials. Likewise, income growth among participant entrepreneurs is anomalous in LEB, with its diminishing returns, but GJ with its *Ak* technology and explicit modeling of information flows provides an explanation. Related also is the diverging income levels between the participant group and the nonparticipant group, not captured in LEB, but reconciled by GJ with its transaction cost.

By contrast, GJ leaves some open questions and anomalies of its own. GJ captures endogenously the overall trend in financial sector participation but cannot explain the relatively sharp upturn in participation after 1986. (The expansion was imposed exogenously in LEB.) GJ is missing the catchup for nonparticipants, no doubt because of missing the wage growth effect, which is picked up well by LEB.

Both models fail to predict the close relationship between the growth patterns of the entrepreneurs in financial sector (the smallest but richest group in the Thai economy) and the aggregate movements of growth and inequality over time. GJ simply does not distinguish occupations. LEB does feature occupations, but the entrepreneurs *in* the financial sector are predicted to have diminishing income growth. Furthermore, the income of entrepreneurs outside the financial sector increased after 1992 even as income growth of entrepreneurs in the financial sector declined. That is, in the data, entrepreneurs seem to face different kinds of technology and shocks, depending whether or not they participate in the financial sector. Part of the decrease in inequality after 1992 is related to the differential income growth between the participant entrepreneurs and nonparticipant entrepreneurs. These phenomena should not be ignored in future work.

Many of these failures seem to be related to insufficient heterogeneity and overly simplistic model structures. A remedy might be to introduce more heterogeneity and additional features of various kinds. That is, it is tempting to think that all one needs to do is to make the models more realistic by adding more kinds of heterogeneity. However, we have learned from the model comparison exercise of this paper that additional heterogeneity *per se* does not necessarily help to improve dynamics or cross-sectional income variation. LEB has more key categories than GJ, and LEB is forced to replicate the financial sector expansion in the data. However, LEB does worse in predicting distributional shapes, subgroup inequality dynamics (e.g., virtually no inequality among non-participants), and income gaps across subgroups. Thus, the compositional effects and convergence/divergence effects are more at odds with the data in LEB than in GJ.

3. LEB MODEL

3.1. Model Economy

We first consider a model of occupational choice, constrained by wealth because of the presumed lack of a credit market as in Lloyd-Ellis and Bernhardt (2000). The economy is populated by a continuum of agents of measure one evolving over discrete time $t = 0, 1, 2 \ldots$ An agent with end-of-period wealth W_t at date t maximizes individual preferences over consumption c_t and wealth carry-over b_{t+1} as represented by the utility function:

$$u(c_t, b_{t+1}) = c_t^{1-\varpi} b_{t+1}^{\varpi},$$

subject to the budget constraint $c_t + b_{t+1} = W_t$.

There are two kinds of production technologies. In traditional sector, everyone earns a safe subsistence return γ of a single-consumption good. In the modern sector, entrepreneurs hire capital k_t and labor l_t at each date t to produce the single-consumption good according to a production function:

$$f(k_t, l_t) = \alpha k_t - \frac{\beta}{2} k_t^2 + \xi l_t - \frac{\rho}{2} l_t^2 + \sigma l_t k_t.$$

Each worker provides a single unit of time and is paid by wage w_t at date t. The cost of capital is determined by its opportunity cost, a constant interest rate of unity tied to a backyard technology. There is a fixed cost of entry into business in the modern sector. That is, the entrepreneur pays an initial setup cost x_t to start up a business. The setup cost represents the inverse of the innate entrepreneurial talent of each agent, and it is assumed to be independent of the wealth level b_t and randomly drawn from a time invariant cumulative distribution:

$$H(x) = mx^2 + (1 - m)x. \qquad \textbf{(1)}$$

The support of x is unit interval $[0, 1]$ and the range of possible values for parameter m is $[-1, 1]$. This class of distributions subsumes the uniform distribution at $m = 0$. As m increases toward 1, the distribution of x becomes more skewed to the right and, hence, efficient entrepreneurs become rare.

In this model, an agent is distinguished by a pair of beginning-of-period characteristics: initial wealth b and randomly drawn entrepreneurial (lack of) talent x, where we suppress the time subscript on these to emphasize the recurrent or stationary aspect. With the above utility function, the optimal rules for consumption and saving will be linear functions of wealth, and so preference maximization is equivalent to end-of-period wealth maximization. Thus, given an equilibrium wage rate w, an agent of type (b, x) chooses his occupation to maximize his total wealth W:

$$
\begin{aligned}
W &= \gamma + b, \quad \text{for subsisters}\\
&= w + b, \quad \text{for wage earners}\\
&= \pi(b, x, w) + b, \quad \text{for entrepreneurs,}
\end{aligned} \tag{2}
$$

where

$$\pi(b, x, w) = \max_{k,l}\{f(k, l) - wl - k - x\} \quad \text{subject to} \tag{3}$$

$$0 \le k \le b - x. \tag{4}$$

Equation (2) suggests that there is a reservation wage level $\underline{w} = \gamma$ below which every potential worker prefers to remain in subsistence sector. Likewise, if the wage rate exceeds that reservation wage, no one remains in subsistence sector. Therefore, the model implies that wage must be $\underline{w} = \gamma$ when the subsistence sector coexists with the modern sector. We allow the subsistence income γ to grow exogenously at the rate of g_γ. As long as two sectors coexist, the demand for labor from the modern sector determines the population proportions of wage earners and subsisters.

The higher is the initial wealth b, the more likely it is that an agent will be an entrepreneur. A potentially efficient, low x, agent may end up being a worker, constrained by low initial wealth b. Given wealth b and market wage w, we can define a marginal agent as one with setup cost $x^m(b, w)$ who is indifferent between being a worker and being an entrepreneur such that $\pi(b, x^m, w) = w$. If the randomly assigned setup cost is higher than this setup cost, the household will be a worker for sure. However, with the constraints on capital demand in (4), the setup cost x cannot exceed the own wealth b either. Therefore, given wage w, the critical setup cost for the marginal agent with wealth b, who is willing to be an entrepreneur, is characterized by

$$z(b, w) = \min[b, x^m(b, w)]. \tag{5}$$

This is the key selection equation, which will be used later in estimating the parameters of the LEB model.

In sum, households of varying talent face imperfect credit markets in financing the establishment of modern business and in expanding the scale of enterprise. Thus, households are constrained by limited wealth on an extensive margin of occupational choice and intensive margin of capital utilized, although both constraints can be alleviated over time. As the distribution of wealth evolves, so does the occupational composition of population and income differentials, generating the dynamics of growth with changing inequality.

3.2. Estimation

3.2.1. Likelihood function. Wealth-constrained occupational choice is the key building block or micro foundation of the LEB model, and the mapping from wealth to occupation itself is stationary, conditional on wage w. Thus, we form a likelihood function of occupational choice as is implied by the theory, and then estimate the key parameters by maximizing the likelihood of the micro data. Because households will at best be indifferent between being wage earners and being subsisters, the crucial occupational choice is binary, between being an entrepreneur or not.

Let y_i denote the binary occupational choice of agent i that assigns 1 for being an entrepreneur and 0 otherwise. Then, given wage w, the probability of being an entrepreneur for agent i with wealth b_i is given by

$$\Pr\{y_i = 1\} = \Pr(x_i \le z(b_i, w)), \tag{6}$$

where z is the critical setup cost function, defined in (5).[5] Then, given the profiles of occupational choice and initial wealth $(y_i, b_i)_{i=1}^n$ of n households in cross-section data, the log likelihood function is written:

$$\log L = \sum_{i=1}^{n} \{y_i \ln[\Pr(x_i \le z(b_i, w))] + (1 - y_i) \ln[1 - \Pr(x_i \le z(b_i, w))]\}, \tag{7}$$

where

$$\Pr(x_i \le z(b_i, w)) = mz(b_i, w)^2 + (1 - m)z(b_i, w), \tag{8}$$

from the time-invariant distribution of random setup cost x, specified in (1).

The critical setup cost function $z(b, w)$ is determined by the optimal profit function in (3) and, hence, by the production technology parameters $(\alpha, \beta, \xi, \rho, \sigma)$. Used also in equation (8) is the parameter m of the talent distribution H. We can thus apply maximum likelihood methods, taking the log likelihood function in (7) to the data on occupational choice in the real Thai economy, to estimate the parameters of technology and distribution of the random talent (setup cost). We may interpret the chosen parameters from this estimation as those that best fit the micro foundation of the LEB model.

Two things are to be mentioned about the estimation. First, because of the quadratic form of technology, the optimal profit function can be written as a reduced-form second-degree polynomial in capital and, hence, only three out of five technology parameters can be identified if we use a single wage. However, by varying the wage over initial time periods in relation to the exogenous parameter of subsistence income γ, we can solve this identification problem. The details of identification are discussed in Appendix A.2. Second, not all remaining parameters can be estimated. The subsistence income level γ and the preference parameter ϖ, the marginal propensity to save, are not directly related to occupational choice and cannot be identified from this estimation. Both ϖ and γ are calibrated below.

3.2.2. Estimates. We estimate and simulate the model using the Thai Socio-Economic Survey (SES), a nationally representative household survey in Thailand for the two decades between 1976 and 1996. The economically active households in the SES data are used. More details of the SES data are described in Appendix A.1.

In order to estimate the mapping between *initial* wealth and *subsequent* occupational choice, as the model suggests, we use only the sample of "young" households whose heads' age is below 30.[6] The choice of cutoff age for "young" households depends on how closely their *current* wealth approximates their *initial* wealth because there is a possibility of wealth accumulation over time, even in the early careers of young households. Thus we compare the cohort age profiles of wealth between the young household group (age < 30) and the rest (age ≥ 30), plotted in Figures A.1 and A.2. We can see that the age profiles of wealth of the young households are literally flat or at least much flatter than those of older ones. Figures A.3 and A.4 compare the age profiles of wealth for entrepreneurs only. Reassuringly, the age profiles of wealth of the young entrepreneurs are flatter than those of older ones, except the latest cohort. Thus, in the data, young entrepreneurs do not accumulate wealth at a high rate so that the current wealth we observe would be close to initial wealth.

The likelihood function in (7) is written for the benchmark LEB model without credit and so we exclude households who participate in the financial sector, to make consistent use of data in estimation. We also use the wage variation only at the initial two years, 1976 and 1981, during which the Thai wage is considered to be close to the reservation wage that grows exogenously in the model.

There is an additional parameter implicitly involved in estimation, the *scale* that converts wealth in the data (in Thai baht unit) into wealth in the LEB model. The choice of the scale is important because the random setup cost x is specified with bounded support [0, 1], and enters in the model in an additive way. We take this scale to be a free parameter and calibrate it as $6 * 10^{-8}$. The way we calibrate the scale will be discussed in the next subsection.

Table 1 reports the estimated parameter values from the maximum likelihood estimation.[7] The average value of log likelihood is -0.4898. The bootstrap

TABLE 1. Estimated LEB parameters

α	β	ξ	ρ	σ	m
1.0011	0.0940	0.0566	0.0033	0	−1
(0.2559)	(0.0103)	(0.0007)	(0.0000)	(0.0001)	(0.0000)

standard errors of the estimates are reported in parentheses.[8] The standard errors are all small and the parameters are fairly precisely estimated.

The LEB model implies an occupational map that partitions the type space (b, x) into four areas: (1) an area of unconstrained subsisters and wageworkers (whose fixed costs are too high, higher than some critical level $x^*(w)$, for them to be entrepreneurs regardless of wealth levels); (2) an area of constrained subsisters and wageworkers (whose fixed costs are lower than $x^*(w)$ but their wealth levels are not high enough to self-finance the fixed costs to be entrepreneurs); (3) an area of constrained entrepreneurs (whose fixed costs are low enough and wealth levels high enough to cover the fixed costs, but not sufficient to finance the unconstrained level of working capital); (4) and an area of unconstrained entrepreneurs (whose wealth levels are sufficient, higher than a critical level $b^*(w)$, to cover both their fixed costs and the unconstrained level of working capital). Let $\widetilde{b}(w)$ be the wealth level below which the wealth constraint binds exactly at the level of setup cost $(x = b)$ and, hence, the capital demand k hits the lower bound at zero. The three parameters $x^*(w)$, $b^*(w)$ and $\widetilde{b}(w)$ determine the shape of the occupational map, conditional on a given wage w.[9] The occupational map at 1976 wage at the above estimates is displayed in Figure 1.

3.3. Calibration

As was mentioned earlier, the wealth scale is a free parameter and estimation is defined conditional on the scale. This is the key parameter that generates a trade-off between cross-sectional estimation and dynamic simulation. We find that the estimated fractions of entrepreneurs in the LEB parameter space are always lower than those in the data, but the higher the scale is, that is, the wealthier the LEB economy becomes, the higher is the likelihood of estimating the cross-sectional occupational choice. But more wealth makes the LEB agents consume their wealth rather than save it and the economy suffers from negative growth initially.[10] The higher the scale, the more negative is this initial negative growth. Although this relationship is not monotone, and eventually the economy starts to grow, as in the data, overall growth for the entire sampling period can be negative when the initial negative growth is too large. So we restrict our search for scale parameter such that this does not happen. The set of estimates reported in Table 1 is the one with the highest likelihood within this range of scales.

The subsistence income γ, its exogenous growth rate g_γ, and the preference parameter ϖ are not related to the occupational choice and they cannot be

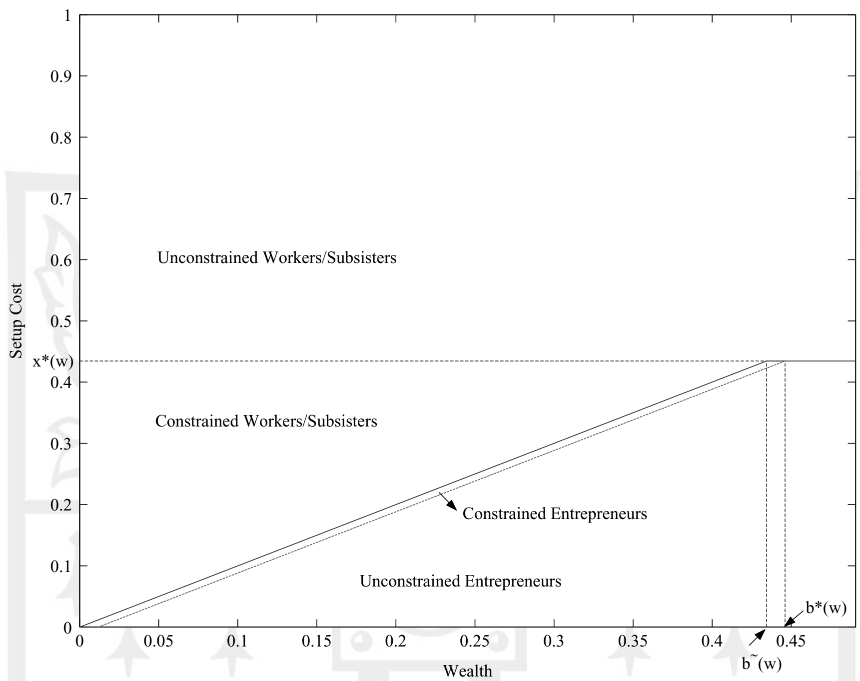

FIGURE 1. Estimated LEB occupation map.

determined from the above estimation. The subsistence income γ is calibrated at 0.012 to match the initial average wage income in 1976, given the above chosen scale. This calibration is equivalent to assuming that the 1976 wage in Thailand is close to the reservation wage. We allow this reservation wage to grow exogenously at the rate of 0.5 percent per year, which matches the annual average growth rate of wage income during the first decade 1976–1986. The wage income in Thailand surged only after 1986 and we consider that the wage growth before 1986 is a result of exogenous growth of reservation wage. The Cobb-Douglas form of preferences implies that ϖ can be interpreted as a savings rate. Thus, we calibrate ϖ at 0.25, matching the average saving rate (with standard error of 0.0204) from the SES during the 1976–1996 period. Thus, all free parameters are calibrated from the same SES data that are used for estimation.

3.4. Evaluation

3.4.1. Aggregate dynamics. The simulated aggregate dynamics paths of LEB are compared with those in the Thai data in Figure 2.[11] The model does capture the overall growth and particularly the accelerated upturn in growth starting 1986 (Figure 2.2). The initial growth rate of the model is higher than the data, but it quickly approaches the low rate of growth in the Thai data, near zero, before the

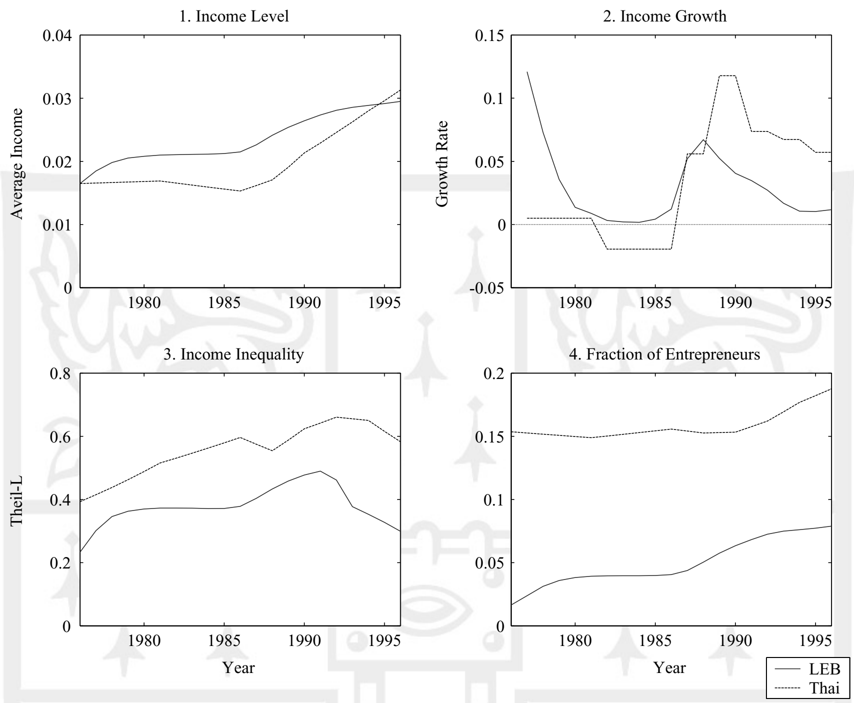

FIGURE 2. LEB aggregate dynamics.

upturn of late 1980s. Inequality dynamics, in particular the overall increase and eventual decrease beginning in early 1990s in Thailand, are also captured, but the predicted level of total inequality in the model is consistently lower than in the data (Figure 2.3). The fraction of entrepreneurs increases both in the model and the data (Figure 2.4). However, the model starts at a lower fraction of entrepreneurs and predicts a noticeable rise during the expansion of the financial sector after 1986, whereas the fraction rises a little later in the data, in 1990.

Evidently, LEB can capture both growth and inequality aggregate dynamics without aggregate shocks, but, as will be shown in the following sections, it does so through endogenous changes in factor prices, that is, wages and interest rates, and through endogenous occupational shifts. However, the exogenously embedded financial expansion is the force behind both the growth dynamics and population dynamics.

3.4.2. Population dynamics. We categorize the population into four subgroups, distinguishing both financial participation and occupation. In the legends in the Figures hereafter, “np” denotes nonparticipants in the financial sector, “p” participants in the financial sector, “e” entrepreneurs, and “ne” nonentrepreneurs. (Nonentrepreneurs include both workers and subsisters, but we will sometimes use

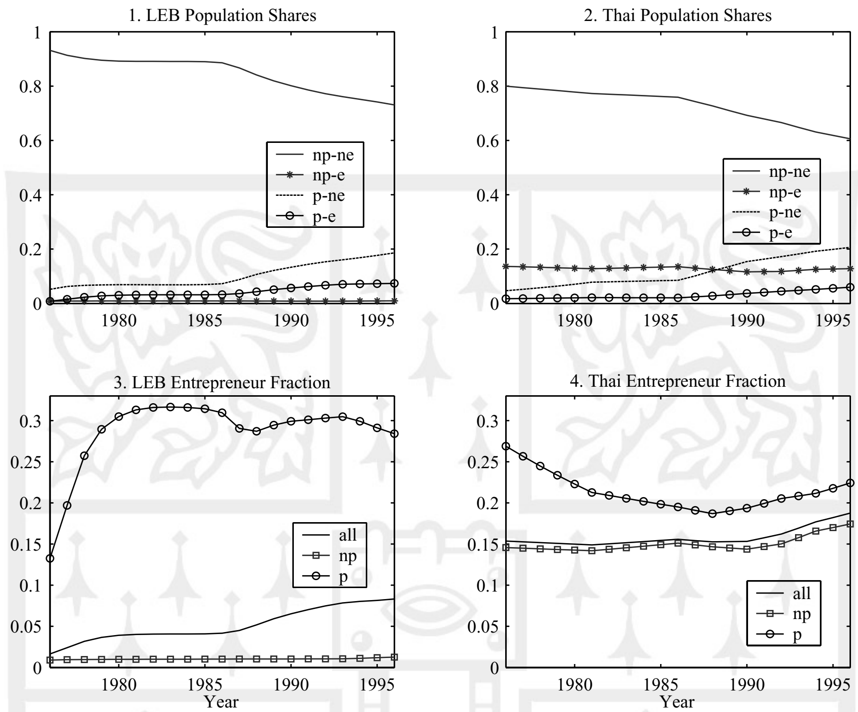

FIGURE 3. LEB population dynamics.

workers and nonentrepreneurs interchangeably because workers and subsisters are all alike in LEB.)

Population shares of the four subgroups are plotted in Figures 3.1 (LEB) and 3.2 (Thai). The directions of compositional change in the model agree with the data. The majority of households are the nonparticipant nonentrepreneurs, but their population share is declining over time in the model, as shown in the data. The population share of nonparticipant entrepreneurs is low and stable over time, as in the data. The population shares of participants of each occupation are increasing in both the model and the data. However, the smallest group are nonparticipant entrepreneurs in the model, whereas the participant entrepreneurs are the smallest group in the data.

The model predicts a larger fraction of entrepreneurs among participants than among nonparticipants, as in the data, although the difference is larger in the model than in the data (Figures 3.3 and 3.4). Except for an initial increase among participants, the fraction of entrepreneurs is more or less stable in the model, among both participants and nonparticipants in the model. Thus, the increase of the economy-wide population share of entrepreneurs is a result of the expansion of financial sector, where there are more entrepreneurs.

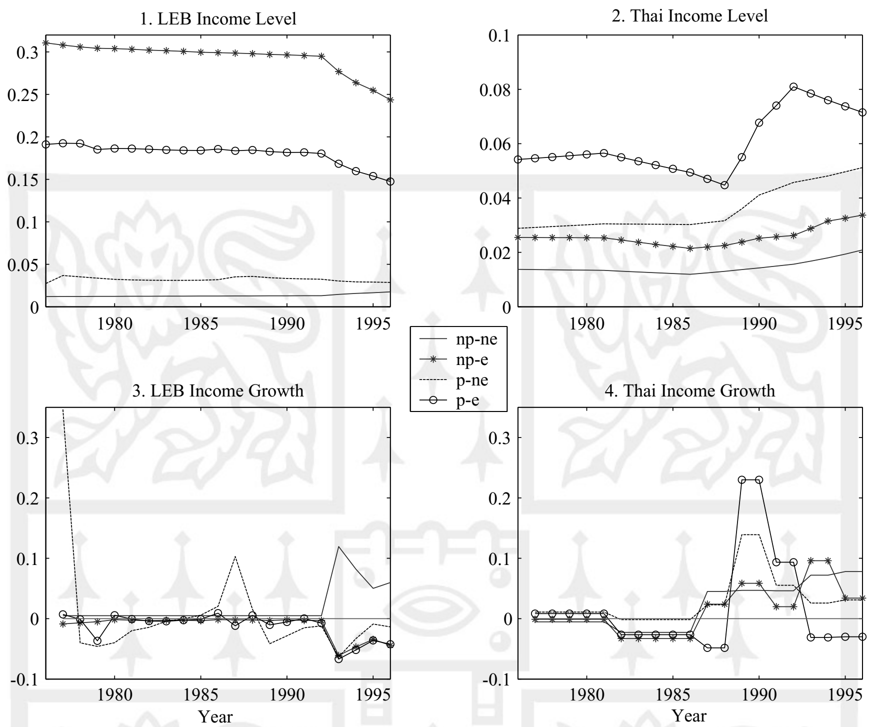

FIGURE 4. LEB subgroup growth dynamics.

3.4.3. Subgroup dynamics. The patterns of subgroup income levels in LEB are displayed in Figure 4.1, juxtaposed with the Thai data in Figure 4.2. Entrepreneurs earn higher income than nonentrepreneurs either within or outside of the financial sector, also true in the data. The richest group in the model are nonparticipant entrepreneurs. However, the richest in the data are *participant entrepreneurs*. In fact, nonparticipant entrepreneurs are poorer than participant workers in the data. Thus, nonparticipant entrepreneurs in the model are too rich. With no access to credit, only the very wealthy can become entrepreneurs to self-finance both the setup cost and the capital. Talented but poor people may not become entrepreneurs if they are outside the credit sector. In contrast, in the credit sector, the poor can be entrepreneurs if their setup cost is low enough. But talented poor people who borrow in the credit sector need to pay the loan back with interest, leaving them poorer in net income than the nonparticipant entrepreneurs who self-finance. (Note that entrepreneurs in both sectors do share the common wage as well as technology.) Thus, on average, entrepreneurial income is higher among nonparticipants than among participants in the model. (The reverse is true in the data.) We may interpret this entrepreneurial income differential of the model as a "rent" due to the imposed factor market structure, that is, a segmented capital market but with an integrated labor market.

In the model, the income of nonparticipant workers grows slowly but steadily but the income of participant workers does not show any trend. In the data, the increasing trend of income of workers is more salient for both participants and nonparticipants. Entrepreneurial income continually decreases over time among both participant and nonparticipant groups in the model, as a result of the diminishing returns to capital in the LEB technology. This decline is accelerated after the wage starts to grow endogenously in 1991. In the data, entrepreneurial incomes also decline for both participants and nonparticipants for the first decade. But after 1986, the income of nonparticipant entrepreneurs steadily *increases*, peaking after 1992. The income of entrepreneurs in the financial sector increases greatly during 1988–1992, then decreases after 1992. Thus, the movements of profits differ between participants and nonparticipants in the data. This suggests a possibility that the participant group and the nonparticipant group may not share a common technology in the data.

Observing the subgroup growth patterns in the data delivers us another interesting point. Note that the period of accelerated income growth of the participant entrepreneurs corresponds to the growth-peak period of aggregate income. Also the point at which the income growth rate of this group begins to decrease corresponds to the turning-point of the aggregate income inequality, from an increasing to a decreasing trend. That is, in the data, the nonlinear dynamics of aggregate income growth and aggregate inequality change are closely related to the growth pattern of *entrepreneurs in the financial sector, the richest subgroup,* although their population share is quite low, going from 2 to 6 percent over time. In LEB, the patterns of entrepreneurial income growth are common between the participants and nonparticipants, that is, the continual decline of profits, because of the common diminishing-return technology. Thus, LEB cannot capture the relationship between the aggregate movement of income growth and inequality and the income growth pattern of the *participant entrepreneurs* that is observed in the data.

Income growth rates co-move in the data across occupation groups within the financial sector (Figure 4.4). Although much less obvious, this co-movement exists in the model, which is related to the movement of the interest rate and hence financial income (Figure 4.3). The co-movements of growth rates across occupation groups among nonparticipants are weak, particularly in the model. In fact, the growth rate of the nonparticipant workers is counter to that of nonparticipant entrepreneurs during the period of endogenous wage growth in the model, and this is also true in the data for the same reason toward the very end of the second decade. However, the model does not capture the co-moving low growth rates among nonparticipants for the first decade.

The entrepreneurial income premium is shown in Figure 5.1, in comparison with the Thai data in Figure 5.2. The income premium of entrepreneurs over wage earners decreases over time for both nonparticipants and participants in the model, as in the data. The eventual decrease in aggregate inequality, shown

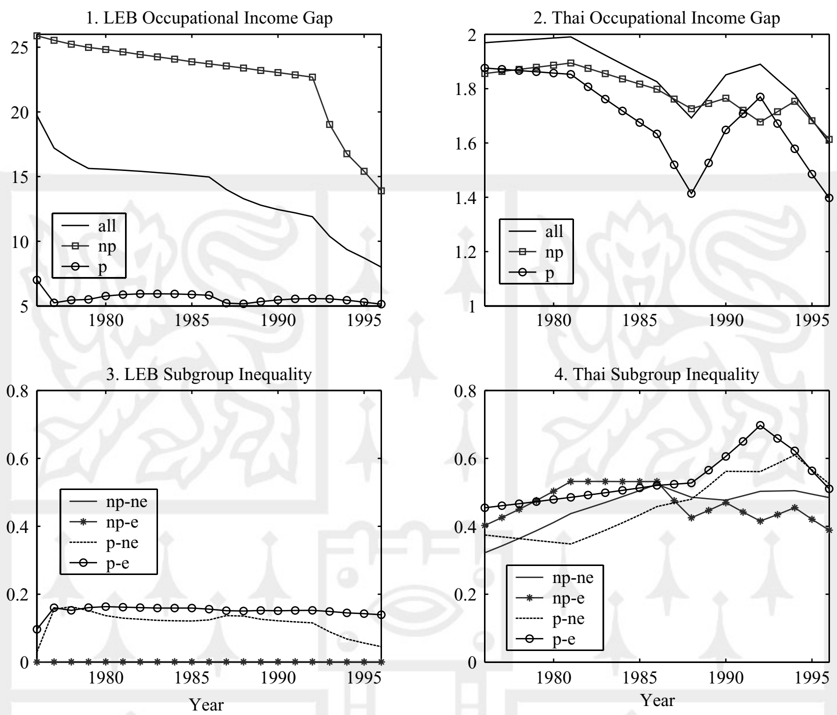

FIGURE 5. LEB subgroup inequality dynamics.

earlier in Figure 2.3, is mainly driven by (endogenous) wage growth and, hence, a decrease in the occupational income gap, as in the data. However, the income gaps are much larger, going from 26 to 13 among nonparticipants and from 7 to 5 among participants in the model than in the data, varying from 1.89 to 1.60 among nonparticipants, and from 1.88 to 1.40 among participants. These occupational income gap changes are mirror images of the above subgroup growth patterns, that is, the continual increase of wage income among the nonparticipant workers, the poorest group, and the continual decrease of profit income among the nonparticipant entrepreneurs, the richest group. This in fact is the source of decreasing aggregate inequality after 1992. Thus, the direction of the occupational income gap change over time agrees with the data, but there is a huge discrepancy in orders of magnitudes.

The model predicts a clear inequality ordering between participants versus non-participants, for each occupation, in Figure 5.3. Inequality levels are much higher among participants than among nonparticipants. In fact, there is literally no inequality among nonparticipant workers, and the inequality among nonparticipant entrepreneurs is virtually nil. The higher level of inequality among participants is due in part to interest income, amplifying wealth differences into income differences, and in part due to more talent variation among entrepreneurs with access to credit. In the data, in Figure 5.4, the inequality-ordering across subgroups is less clear, except that income inequality of entrepreneurs in the financial sector is higher than workers in the financial sector. The Thai data show co-movements of inequality levels across occupation groups in the financial sector but this is weak in the model. Finally, the model predicts much lower subgroup inequality levels for all four groups than in the data. Income variation within every subgroup is too small in the model relative to the data.

3.4.4. Decomposition formulae. Aggregate dynamics are generated from the above population dynamics and subgroup dynamics. We decompose the aggregate growth of mean income and income inequality into the contributions of those underlying components according to the following formulae.

The aggregate mean income μ is a sum of subgroup mean income μ^k's, weighted by subgroup population shares p^k's:

$$\mu = \sum_{k=1}^{K} p^k \mu^k.$$

The change in mean income is thus decomposed into two parts, one from the changes in population shares Δp^k's, and the other from growth within subgroups $\Delta\mu^k$'s:

$$\Delta\mu = \sum_{k=1}^{K} \overline{p^k}\,\Delta\mu^k + \sum_{k=1}^{K} \overline{\mu^k}\,\Delta p^k, \tag{9}$$

where Δ denotes the difference over time and the upper bar the average over time. This is simply a discrete version of product rule, which can be applied to any additive indices. We will use the decomposition formula in terms of overall growth *rate*, by measuring changes between the beginning and ending points for two decades, dividing all terms in (9) by the initial mean income level.

Theil-L entropy index I, our measure of income inequality, is also additively decomposable into *within-group inequality* WI and *across-group inequality* AI as follows:

$$\begin{gathered} I = WI + AI, \\ WI = \sum_{k=1}^{K} p^k I^k, \quad AI = \sum_{k=1}^{K} p^k \log\left(\frac{\mu}{\mu^k}\right), \end{gathered} \tag{10}$$

where I^k denotes the inequality within subgroup k. The within-group inequality WI is a sum of the subgroup inequality I^k's weighted by the population shares of subgroups. The across-group inequality AI is a sum of log inverse of relative incomes, again weighted by the population shares of subgroups.

Because of the additive nature of the Theil-L index, we can also apply this discrete product rule to the inequality change over time as follows:

$$\Delta I = \Delta WI + \Delta AI, \qquad \textbf{(11)}$$

$$\Delta WI = \sum_k \overline{p^k}\Delta I^k + \sum_k \overline{I^k}\Delta p^k,$$

$$\Delta AI \doteqdot \sum_k \left[\overline{\left(\frac{p^k\mu^k}{\mu}\right)} - \overline{p^k}\right]\Delta \log \mu^k + \sum_k \left[\overline{\left(\frac{\mu^k}{\mu}\right)} - \overline{\log\left(\frac{\mu^k}{\mu}\right)}\right]\Delta p^k. \quad \textbf{(12)}$$

The change in total inequality is decomposed into the change in within-group inequality ΔWI and the change in across-group inequality ΔAI.[12] The within-group inequality change ΔWI is further decomposed into the change in subgroup inequality and the change in population composition as in (12). The across-group inequality change ΔAI is decomposed into two components as well: the change in relative income gaps across subgroups, the first term in (12), and the change as a result of the population composition changes, the second term in (12). Note that $\Delta \log \mu^k$ approximates the growth rate of average income of subgroup k; hence, the first term in (12) captures the inequality change because of the differential growth rates across subgroups weighted appropriately. When a higher-income group grows faster than a lower-income group, the income gap between the two diverges and the across-group inequality increases (or vice versa). We thus call this term *divergence (or convergence) effect.* We will use the decomposition formulae by normalizing the terms in (12) and (12) by the initial inequality level.

3.4.5. Decomposition results. The two-decade growth rate of mean income is decomposed into the contributions of subgroup growth and compositional growth for both the LEB model and the Thai data in Table 2, using the above formula in (9). We partition the population: by occupation category only, by financial participation only, and then by joint categories distinguishing both occupation and financial participation.

The model almost matches the overall income growth rate (0.869 for LEB and 0.899 for Thailand). However, growth comes mainly from occupational shifts in the model (at the rate of 0.754), but this composition effect on growth is small in the data (at the rate of 0.032). Partitioning the population by financial participation and ignoring the difference in occupation, the composition effect from the (exogenous) expansion of intermediation contributes substantially to growth in the model (at the rate of 0.456) as in the data (at the rate of 0.319). Distinguishing the population by both occupation and financial participation, we observe in the model a magnitude of composition effect similar to the case when only the occupation is distinguished. This implies that the huge dominance of the composition effect on growth in the model is mainly due to the enormous *occupational* income gaps (varying from 26 to 13 among nonparticipants and from 7 to 5 among participants) rather than the income gap between financial participants and nonparticipants.

TABLE 2. Decomposition of aggregate income growth in LEB

	Subgroup	Composition	Total
By Occupation			
Thailand	0.867	0.032	0.899
LEB	0.115	0.754	0.869
By Financial Participation			
Thailand	0.580	0.319	0.899
LEB	0.413	0.456	0.869
By Joint Category			
Thailand	0.573	0.326	0.899
LEB	0.141	0.728	0.869

The total inequality level can be decomposed into within-group inequality and across-group inequality as in equation (10). This is displayed in Figure 6, for both LEB and Thailand, taking the population partition by joint categories. This suggests that across-group inequality is the main component of total inequality in the model while within-group inequality is the main one in the data.

The two-decade growth rate of total inequality is decomposed for both the model and the data in Table 3, using equations (12) and (12).[13] The model predicts an overall increase in inequality at 0.338 but this is less than in the data at 0.483.

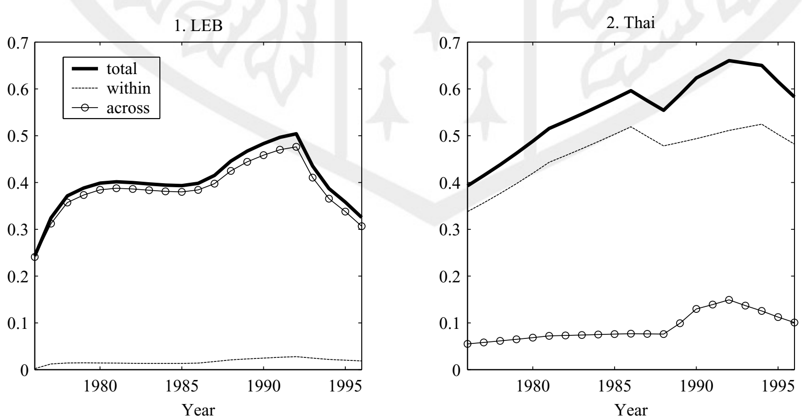

FIGURE 6. Within vs. across inequality decomposition.

TABLE 3. Decomposition of aggregate inequality change in LEB

By Occupation

	Within-Group		Across-Group		
	Subgroup	Composition	Income Gap	Composition	Total
Thailand	0.524	0.001	−0.051	0.010	0.483
LEB	0.042	0.022	−1.056	1.881	0.338

By Financial Participation

	Within-Group		Across-Group		
	Subgroup	Composition	Income Gap	Composition	Total
Thailand	0.304	0.032	0.015	0.133	0.483
LEB	−0.177	0.189	−0.066	0.439	0.338

By Joint Category

	Within-Group		Across-Group		
	Subgroup	Composition	Income Gap	Composition	Total
Thailand	0.340	0.028	−0.003	0.120	0.483
LEB	0.015	0.053	−0.750	1.371	0.338

Distinguishing the population by occupation only, the model predicts an increase in subgroup inequality, a decrease in inequality through a converging occupational income gap, and an increase in inequality through the two composition effects. The directions of all these effects on inequality change in the model are consistent with the data. However, the orders of magnitudes of all these effects are quite different from the data. In the model, the subgroup inequality change is too small, and the convergence and composition effects are much too big.

Distinguishing the population by financial participation only, the model delivers a significant composition effect of financial expansion on across-group inequality, as in the data. However, subgroup inequality levels among both participants and non-participants decrease in the model, different from the data. This is because of the decrease in occupational income gap within each sector. Also, the model predicts convergence in income levels between participants and nonparticipants, but we observe divergence in the data. Thus, exogenous incorporation of financial expansion helps to explain the composition effects on inequality change (and income growth as well) but creates anomalies in other dimensions. We will see if endogenizing the financial participation decision can remove these anomalies in the LEB model. Distinguishing by both characteristics, the features of decomposition are similar to the decomposition by occupation only, but the difference in magnitudes between the model and the data becomes smaller.

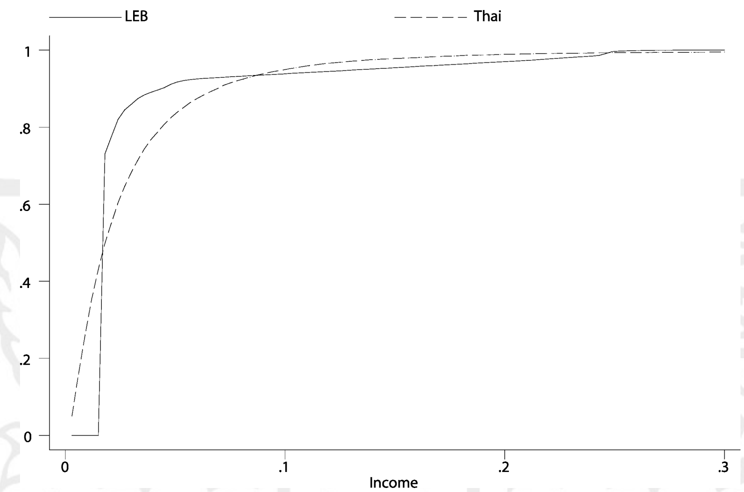

FIGURE 7. CDF of 1996 income distribution.

3.4.6. End-of-sample-period income distribution. The cumulative distribution functions of the income distributions at the end of sample period, 1996, are plotted for both the LEB model and the Thai data in Figure 7. The main discrepancy comes from the lower tail of the distribution. There is a spike at the low end of income distribution due to the common wage in the model, whereas there is much more income variation within the lower tail in the data. In the upper tail of the distribution, there is a slight bimodality in the model due to the income gap between nonparticipant entrepreneurs and participant entrepreneurs, and there are no extremely rich people that are present in the data. Thus, the model does not capture the income variation at both lower and upper tails in the data.

We formally test the goodness of fit of the end-of-sample-period income distribution of LEB relative to the Thai data. The income distribution in the data is not necessarily close to some *a priori* parametric form. The income distributions from the model are endogenously determined, evolving over time and would be distorted by imposing parametric forms on them. Thus, we compare distributional shapes between the model and the data in a *nonparametric* way, using the Kolmogorov-Smirnov (KS) statistic:

$$KS = \sqrt{mn/(m+n)} \sup_{-\infty<y<\infty} |F_n(y) - G_m(y)|, \qquad (13)$$

where G_m and F_n denote the empirical distribution functions from the model and the data, respectively, and m and n denote the sample size of the empirical distributions from the model and the data, respectively. The limiting distribution of this statistic is described in Smirnov (1948). The KS statistic for the LEB model is 3.04 and the corresponding p-value is less than 0.0000, strongly rejecting similarity of income distributions between the LEB model and the Thai data.[14] This rejection is obviously due to the spike at the low end of the LEB distribution.

3.5. Sensitivity Analysis

We perform a sensitivity analysis for the LEB model, to check the robustness of our evaluation results. For the *estimated* parameters such as technology parameters $(\alpha, \beta, \xi, \rho, \sigma)$ and talent distribution parameter m, an obvious concern would be sampling errors around the point estimates. However, the bootstrap standard errors of these estimates are small, virtually zero. In fact, varying the estimated parameters within the range of one or two standard-error bounds does not change the simulation results. Thus we further vary the parameter values within a 10-percent-deviation range.

The simulated dynamics for both growth and inequality turn out to be robust to parameters α, β, and σ, but sensitive to ξ and ρ.[15] Explicit consideration of the profit function of LEB helps us to understand why. Given the quadratic production function, the profit function can be written

$$\pi = C_0(w) + C_1(w)k + C_2k^2 - x, \tag{14}$$

where

$$C_0(w) = \frac{(\xi - w)^2}{2\rho}, \tag{15}$$

$$C_1(w) = \alpha - 1 + \frac{\sigma(\xi - w)}{\rho}, \tag{16}$$

$$C_2 = \frac{1}{2}\left(\frac{\sigma^2}{\rho} - \beta\right). \tag{17}$$

Thus, three coefficients, $C_0(w)$, $C_1(w)$, and C_2, determine the dynamics of profit growth in relation to wage w. These also determine the three occupational map parameters, $b^*(w)$, $x^*(w)$, and $\widetilde{b}(w)$[16] :

$$\widetilde{b}(w) = C_0(w) - w, \tag{18}$$

$$x^*(w) = \widetilde{b}(w) - \frac{C_1(w)^2}{4C_2}, \tag{19}$$

$$b^*(w) = x^*(w) - \frac{C_1(w)}{2C_2}. \tag{20}$$

At our estimate of σ at zero, $C_1(= \alpha - 1)$ and $C_2(= -\frac{2}{\beta})$ are time-invariant and independent from the wage. Thus changes in α and β affect the shape of the profit function, and subsequently the income dynamics, but not in relation to the wage evolution. Changes in α and β affect x^* and b^*, again not in relation to the wage, with σ at zero. Thus, these changes do not shift the occupational map over time.

Varying σ away from zero, a change in α could affect income and occupational choice in relation to wage via $C_1(w)$. However, a range of σ of [0, 0.013] turns out to be not wide enough to generate any significant changes. Figure A.7 shows that an increase in α to 1.3 (maximum allowable value) reduces the annual average income growth rate from 3.18% to 2.95%, and also reduces the annual average rate of increase in inequality from 1.47% to 0.80%. Occupation transition dynamics remain virtually the same.

By contrast, ξ and ρ can directly affect both income dynamics and occupational choice via $C_0(w)$ in relation to wage although σ is near zero. An increase in ξ implies an increase in the intercept term $C_0(w)$ of the profit function. It also implies an increase in marginal productivity of labor by constant term ($MPL = \xi - \rho l$ with σ at zero). This makes the modern business more profitable and draws more savings into the more productive modern business sector. Thus income growth becomes faster. However, its implication for inequality is a bit complicated. When wage is set at the reservation level, the benefits from an increase in labor productivity belong to entrepreneurs but not to workers, and inequality rises. Once the wage starts to grow endogenously, the increase in marginal productivity from an increase in ξ benefits workers. But because both ξ and w increase, the net effect on $C_0(w)$ and on profits is not certain. An increase in ρ plays a similar role to a decrease in ξ. Figure A.8 displays that an increase in ξ by 10 percent to 0.0623 increases the annual average income growth rate from 3.18% to 3.49%, but makes income inequality decline so fast in the end that the overall annual average rate of inequality change becomes negative, to −2.8%. However, the *patterns* of dynamics still remain the same: income grows slowly for the first decade and then surges after the financial expansion; income inequality follows an inverted-U shaped path, and the declining inequality is as a result of the endogenous wage growth. Occupation transition dynamics are again robust to this change.

An increase in m decreases the population of talented people and income growth becomes slower. This also reduces the income gap between entrepreneurs and wage earners, hence the inequality level decreases. An increase in γ makes the economy richer and reduces the occupational income gap. But this is a level effect. In contrast, an increase in g_γ has no level effects but does have exogenous growth effects. An increase in ω induces higher saving. This makes wealth accumulation faster and the occupational transition easier for the constrained workers. Thus income grows faster and inequality starts to decline earlier. However, the orders of magnitude of the changes in aggregate dynamics from perturbing parameters m, γ, g_γ, and ω within 10-percent-deviation bands are small.

4. GJ MODEL

4.1. Model Economy

In Greenwood and Jovanovic (1990), hereafter denoted by GJ, growth and the evolution of the income distribution are related explicitly to financial deepening, that is, increasing participation in the financial sector. There exists a fixed entry cost that endogenously constrains the participation decision.

Consider an economy with a continuum of agents on the unit interval [0,1. They live for an infinite, discrete number of periods $t = 0, 1, 2 \cdots$. For every agent j, there are two technologies available that can convert the capital investment i_{jt} at date t into income $y_{j,t+1}$ at next date $t+1$. One technology yields a safe but relatively low rate of return δ per unit capital and the other gives a risky rate of return $(\zeta_{t+1} + \epsilon_{j,t+1})$ with higher expected value, where ζ_{t+1} represents a common aggregate shock and $\epsilon_{j,t+1}$ an idiosyncratic shock specific to agent j. The aggregate shock ζ_{t+1} is governed by a time-invariant uniform distribution with support $[\underline{\zeta}, \overline{\zeta}]$, and the idiosyncratic shock $\epsilon_{j,t+1}$ is governed by a time-invariant uniform distribution with support $[-\overline{\epsilon}, \overline{\epsilon}]$ with $E(\epsilon_{j,t+1}) = 0$. Let $\eta_{j,t+1} = \zeta_{t+1} + \epsilon_{j,t+1}$ be the composite shock and Ψ^{η} be its cumulative distribution function. GJ assume that the lower bound for the composite shock is positive, i.e., $\underline{\zeta} - \overline{\epsilon} > 0$.

Each agent j decides on running either technologies, with portfolio share ϕ_{jt} for the risky one, at date t so that the next period beginning-of-period wealth $k_{j,t+1}$ can be written such as

$$k_{j,t+1} = [\phi_{jt}\eta_{j,t+1} + (1-\phi_{jt})\delta]i_{jt}. \qquad \textbf{(21)}$$

He allocates his beginning-of-period wealth k_{jt} into current consumption c_{jt} and capital investment i_{jt}, namely, $k_{jt} = c_{jt} + i_{jt}$. The objective is then to maximize the discounted life-time utility stream:

$$E\sum_{t=0}^{\infty}\beta^t\frac{c_{jt}^{1-\sigma}}{1-\sigma},$$

subject to the sequence of resource constraints of $k_{jt} = c_{jt} + i_{jt}$ and the law of motion in (21).[17] Agents are heterogeneous in their wealth levels in each period for two reasons: first, the initial endowment k_0 at date 0 differs across agents, distributed under a cumulative distribution function Λ_0. Second, the history of realizations of random shock up to date $t\{\epsilon_{j,s}\}_{s=0}^{t}$ differs across agents j's.

Other than physical production, there is another "technology" available, namely, *financial intermediation*. An intermediary can run a countable number of trials for the risky technology and get advanced information on next period's return to the risky project. Then, the intermediary invests in the risky project only if this return exceeds the safe return δ. Furthermore, the intermediary can diversify the idiosyncratic shocks $\epsilon_{j,t+1}$ by pooling participants' returns. It can pay back at date $t+1$ a promised return $r(\zeta_{t+1})$, to be spelled out below in (25), per unit of capital invested at time t contingent on the realized aggregate shock ζ_{t+1}. Therefore, every agent has an incentive to join the coalition of financial intermediaries.

There are key restrictions on the parameter space to make the above economy work properly.[18] In order to ensure the benefits of intermediation and the incentive to invest positive amount in the risky production every period, we need to assume the following condition:

$$E\{r(\zeta_{t+1})\} > E\{\zeta_{t+1}\} > \delta. \tag{22}$$

To avoid the economy from shrinking to negative infinity, we need:

$$\delta > 1/\beta. \tag{23}$$

With the linear production technology, unbounded growth is possible in this model, and in order to prevent the economy from exploding to infinity in utility terms, we need

$$\beta E\{r(\zeta_{t+1})^{1-\sigma}\} < 1. \tag{24}$$

These intermediary trading arrangements are costly, as in Townsend (1978). There is an initial fixed cost α of admitting each participant into the financial coalition and a variable cost of $(1-\gamma)$ in proportion to the amount of funds each agent invests in the coalition. Thus, the intermediary charges a lump sum entry fee q for each participant in exchange for the rights to operate an individual's project. The zero-profit condition for the intermediary implies

$$r(\zeta_{t+1}) = \gamma \max\{\delta, \zeta_{t+1}\}, \tag{25}$$

$$q = \alpha. \tag{26}$$

Given these entry and proportional fees, not everyone immediately joins the financial system. Only agents whose wealth levels exceed some critical level are willing to join. That is, the choice of participation in the financial sector is constrained by wealth.

The decision making of households can be characterized by a pair of value functions: v^0, the value function of nonparticipants, and v^1, the value function of participants. An agent j with wealth k_{jt} at date t who currently is not in the intermediation coalition chooses the total investment i_{jt} and the portfolio ϕ_{jt} between safe and risky projects according to the following functional equation:

$$\begin{aligned} v^0(k_{jt}) = \max_{i_{jt},\phi_{jt}} & \big\{u(k_{jt} - i_{jt}) \\ & + \beta E_{\eta_{j,t+1}} \max\left[v^0(k_{j,t+1}), v^1(k_{j,t+1} - q)\right]\big\} \quad \text{subject to (21)}, \end{aligned} \tag{27}$$

where $E_{\eta_{j,t+1}}$ is the expectation with respect to the composite shock $\eta_{j,t+1}$. An agent who is already in the financial system decides only on total investment by solving the following functional equation:

$$\begin{aligned} v^1(k_{jt}) = \max_{i_{jt}} & \big\{u(k_{jt} - i_{jt}) \\ & + \beta E_{\zeta_{t+1}} v^1(k_{j,t+1})\big\} \quad \text{subject to } k_{j,t+1} = r(\zeta_{t+1}) i_{jt}. \end{aligned} \tag{28}$$

Note that the expectation operator in the participant's value function is taken only with respect to the aggregate shock ζ_{t+1} because the idiosyncratic shock is diversified away. Also note that participants do not choose the portfolio between safe and risky investments because this decision is delegated to the intermediary who has advanced information on the aggregate shock. There is no value function for non-participants v^0 in (28) because $v^1(k) > v^0(k)$ for every k, so once an agent enters the intermediated sector, he will never exit.

In sum, households face wealth constraints in their decisions to undertake costly entry into the financial system itself. Participation in financial intermediaries provides the benefits of enhanced risk sharing and advanced information. As economy-wide wealth shifts to the right, more households gain access to financial intermediaries, and this changes the composition of income-status groups and the income gap across them, which in turn over time affects growth and inequality dynamics.

4.2. Estimation

4.2.1. Likelihood function. Financial participation constrained by wealth is the key micro foundation of the GJ model. We form a likelihood function for the financial participation decision using the pair of dynamic programs in (27) and (28). Let d_{jt} denote the participation decision of agent j at date t, which assigns 1 if agent j decides to participate in the financial sector, and 0 otherwise:

$$\begin{aligned} d_{jt} &= 1, \quad \text{if } v^1(k_{jt} - q) \geq v^0(k_{jt}) \\ &= 0, \quad \text{if } v^1(k_{jt} - q) < v^0(k_{jt}). \end{aligned} \tag{29}$$

Townsend and Ueda (2006) show that there exists a unique critical value k^* such that the participation decision in (29) is equivalent to

$$\begin{aligned} d_{jt} &= 1, \quad \text{if } k_{jt} \geq k^* \\ &= 0, \quad \text{if } k_{jt} < k^*. \end{aligned} \tag{30}$$

There is no closed-form solution for k^* because there are no analytic solutions to the dynamic program in (27). However, from the formulation of the dynamic programs in (27) and (28), it is clear that k^* is a function of the underlying parameters of the GJ model, $\theta^{GJ} = (q, \delta, \beta, \sigma, \gamma, \underline{\zeta}, \overline{\zeta}, \epsilon)$,

$$k^* = k^*(\theta^{GJ}).$$

From the recursive nature of the above dynamic programming problems, the policy functions for portfolio and investment are time-invariant and depend on current wealth and the underlying parameters θ^{GJ}, that is, $\phi_{jt} = \phi(k_{jt}, \theta^{GJ})$ and $i_{j,t} = i(k_{jt}, \theta^{GJ})$. Recalling that the law of motion for those outside the financial system is given by (21), a previous non-participant enters the financial sector today at t only if

$$k_{jt} = \left[\phi\left(k_{j,t-1}, \theta^{GJ}\right)\eta_{jt} + \left(1 - \phi\left(k_{j,t-1}, \theta^{GJ}\right)\right)\delta\right] i\left(k_{j,t-1}, \theta^{GJ}\right) \geq k^*(\theta^{GJ}).$$

That is, the participation decision of a previously non-participating agent j at date t with wealth $k_{j,t-1}$ can be rewritten as

$$\begin{aligned} d_{jt} &= 1, \quad \textit{if } \eta_{jt} \geq \eta^*\left(k_{j,t-1}, \theta^{GJ}\right) \\ &= 0, \quad \textit{if } \eta_{jt} < \eta^*\left(k_{j,t-1}, \theta^{GJ}\right), \end{aligned} \tag{31}$$

where:

$$\eta^*\left(k_{j,t-1}, \theta^{GJ}\right) \equiv \frac{1}{\phi\left(k_{j,t-1}, \theta^{GJ}\right)} \left[\frac{k^*(\theta^{GJ})}{i\left(k_{j,t-1}, \theta^{GJ}\right)} - \left(1 - \phi\left(k_{j,t-1}, \theta^{GJ}\right)\right)\delta\right]. \tag{32}$$

The participation decision in (31) is stationary for a given wealth level because the composite technology shock η_{jt} is drawn from the time-invariant distribution. Thus once we solve the pair of functional equations in (27) and (28) to get k^* and the time-invariant policy functions ϕ and i, we can form a likelihood function in terms of model parameters θ^{GJ}.

In forming the likelihood function, the unobservable aggregate shock ζ_t generates cross-sectional dependence over the individuals at a given date t. Thus, we first consider a *conditional* likelihood function $L_t(\theta^{GJ}, \zeta_t)$ for a given aggregate shock ζ_t and then integrate the aggregate shock out to form an unconditional likelihood function, as follows.

Given a series of serially independent aggregate shocks $(\zeta_t)_{t=1}^T$, the likelihood function can be factorized into marginal likelihoods:

$$L\left(\theta^{GJ}, (\zeta_t)_{t=1}^T\right) = \prod_{t=1}^{T} L_t\left(\theta^{GJ}, \zeta_t\right). \tag{33}$$

Because the composite shock $\eta_{jt} = \epsilon_{jt} + \zeta_t$ is *i.i.d. conditional on* ζ_t, a conditional likelihood function $L_t(\theta^{GJ}, \zeta_t)$ at date t is given by:

$$L_t\left(\theta^{GJ}, \zeta_t\right) = \prod_{j=1}^{n_t} [1 - \Pr(\epsilon_{jt} \leq \eta^*_{jt} - \zeta_t)]^{d_{jt}} [\Pr(\epsilon_{jt} \leq \eta^*_{jt} - \zeta_t]^{1-d_{jt}}, \tag{34}$$

where $\eta^*_{jt} = \eta^*(k_{j,t-1}, \theta^{GJ})$ in (32).

Combining the equations (33) and (34), given the data $((k_{j,t-1}, d_{jt})_{j=1}^{n_t})_{t=1}^T$, the conditional log likelihood is written as

$$\ln L\left(\theta^{GJ}, (\zeta_t)_{t=1}^T\right) \tag{35}$$

$$= \sum_{t=1}^{T} \sum_{j=1}^{n_t} \{d_{jt} \ln[1 - \Pr(\epsilon_{jt} \leq \eta^*_{jt} - \zeta_t)] + (1 - d_{jt}) \ln[\Pr(\epsilon_{jt} \leq \eta^*_{jt} - \zeta_t]\}, \tag{36}$$

We integrate the aggregate shocks out by taking expectations in (36) with respect to $\zeta = (\zeta_t)_{t=1}^{T}$:

$$\ln L(\theta^{GJ}) = E_{\zeta} \ln L\big(\theta^{GJ}, (\zeta_t)_{t=1}^{T}\big) \tag{37}$$

$$= \sum_{t=1}^{T} \sum_{j=1}^{n_t} E_{\zeta_t} A_{jt}(\zeta_t), \tag{38}$$

where

$$\begin{aligned} A_{jt}(\zeta_t) &= d_{jt} \ln\left[\frac{1}{2} - \frac{\eta^*_{jt} - \zeta_t}{2\overline{\epsilon}}\right] \\ &\quad + (1 - d_{jt}) \ln\left[\frac{1}{2} + \frac{\eta^*_{jt} - \zeta_t}{2\overline{\epsilon}}\right], \quad \mathit{if} -\overline{\epsilon} \le \eta^*_{jt} - \zeta_t \le \overline{\epsilon} \\ &= (1 - d_{jt}) * (-\infty), \quad \mathit{if}\, \eta^*_{jt} - \zeta_t \le -\overline{\epsilon} \\ &= d_{jt} * (-\infty), \quad \mathit{if}\, \eta^*_{jt} - \zeta_t \ge \overline{\epsilon}. \end{aligned} \tag{39}$$

The $A_{jt}(\zeta_t)$ comes from the uniform distribution of ϵ_{jt}.[19] These terms are numerically integrated with respect to ζ_t according to the uniform distribution over $[\underline{\zeta}, \overline{\zeta}]$ to get $E_{\zeta_t} A_{jt}(\zeta_t)$ in (38). We choose the parameter vector θ^{GJ} that maximizes the log likelihood function in (38), satisfying the restrictions on parameter space given in (22), (23), and (24).

In GJ, the scale parameter matters as in LEB. We choose the GJ scale of wealth by matching the critical wealth level k^* with the wealth percentile $\widehat{k}$ in the data such that the implied GJ participation rate in the financial sector matches the 1976 participation rate in the data.[20] That is, wealth in the Thai data is converted into wealth in the GJ model using the following scale:

$$\text{scale}^{GJ} = \frac{k^*(\theta^{GJ})}{\widehat{k}}. \tag{40}$$

Thus, we compute k^* before estimation to get the scale and then use the scaled wealth of the data in the likelihood function. Thus, we implicitly estimate the GJ scale parameter as well.

4.2.2. Estimates. We again use only the young-household sample (but for all available years) for estimation because the GJ likelihood function maps the *initial* wealth into the *subsequent* participation decision. Thus, we need to restrict our sample for estimation to the households whose current wealth approximates previous wealth. Here we again rely on the evidence on cohort age profiles of wealth (Figures A.5 and A.6). The age profiles of wealth of young participants are still flatter than those of the older ones, except the latest three cohorts.

The estimates from the MLE are reported in Table 4 with bootstrap standard errors in parenthesis. The average value of log likelihood is -0.7116. The value

TABLE 4. Estimated GJ parameters

q	γ	β	σ	δ	$\underline{\zeta}$	$\overline{\zeta}$	$\overline{\epsilon}$
0.5021	1	0.9627	0.9946	1.0479	1.0470	1.1905	0.9954
(0.0482)	(0.0000)	(0.0061)	(0.0926)	(0.0064)	(0.0451)	(0.0514)	(0.0355)

functions v^0 and v^1 at these estimates are plotted in Figure 8, which shows the unique critical wealth level k^* of 1.0292, the crossing point of $v^0(k)$ and $v^1(k-q)$, which partitions the population into nonparticipant and participant groups. Note that the estimated fixed cost parameter q at 0.5021 is about half of this critical wealth. The estimate of γ at one (the most robust estimate) implies no variable cost. Thus, the fixed cost, not the variable cost, plays an important role in intermediation. The estimate of discount factor β at 0.9627 belongs to the range of values that are often adopted in the business cycle literature. It is interesting to note that the estimated relative risk aversion parameter is very close to one, that is, the case of log utility function as in the original GJ paper. The estimates of δ, $\underline{\zeta}$, and $\overline{\zeta}$ imply that the rates of return are 5% to safe investment and 12% to risky investment. The estimated range of idiosyncratic shock [−0.9954, 0.9954] is wide enough that

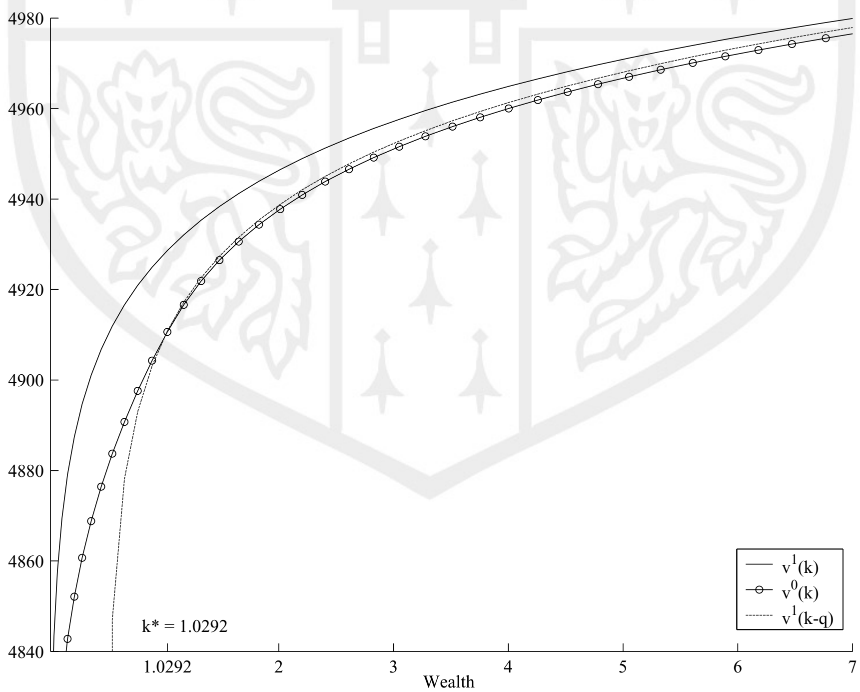

FIGURE 8. GJ value functions.

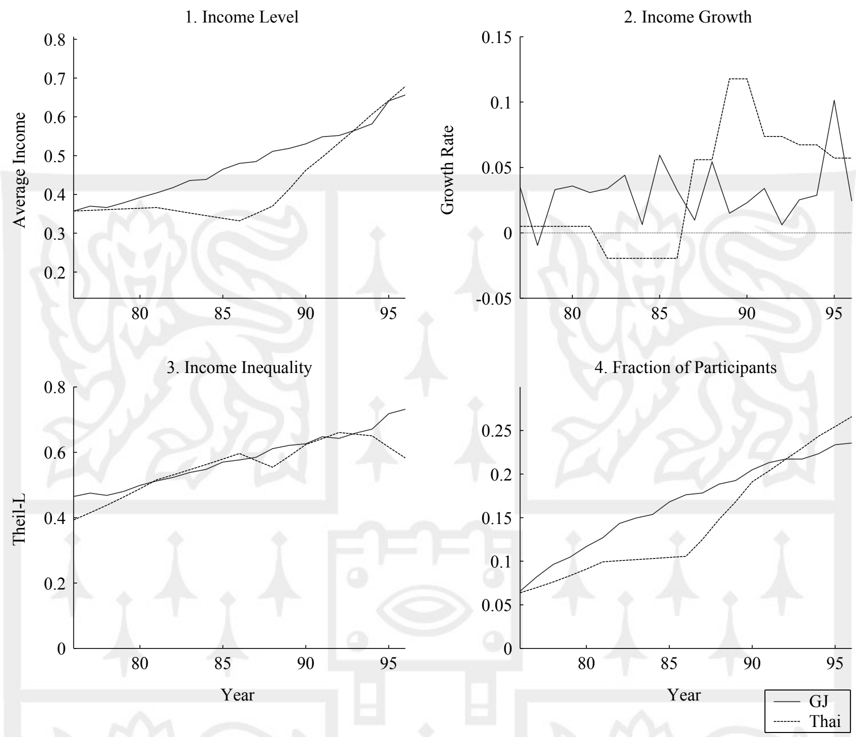

FIGURE 9. GJ aggregate dynamics.

some nonparticipants with low wealth can pay the fixed entry cost 0.5021 by a single lucky draw of idiosyncratic shock.

4.3. Evaluation

4.3.1. Aggregate dynamics. Figure 9 displays the aggregate dynamics of GJ in comparison with the Thai data. The model captures quite well the levels and overall increase in income inequality and to a lesser extent overall income growth.[21] However, the trends are more or less linear in the model and GJ simulation captures neither the initially slow and then accelerated upturn of income growth in early 1990s nor the eventual downturn of income inequality after 1992 in Thailand. The fraction of participants in the financial sector increases both in the model and the data at similar orders of magnitudes. Again, however, the model predicts a linear trend in financial expansion, whereas the data show a clear nonlinear expansion, whereas the substantial acceleration after 1986.

4.3.2. Subgroup dynamics. Figure 10 shows the income dynamics of participant and nonparticipant groups, in comparison with those of Thailand. The growth rate of income of participants is almost always higher than that of nonparticipants

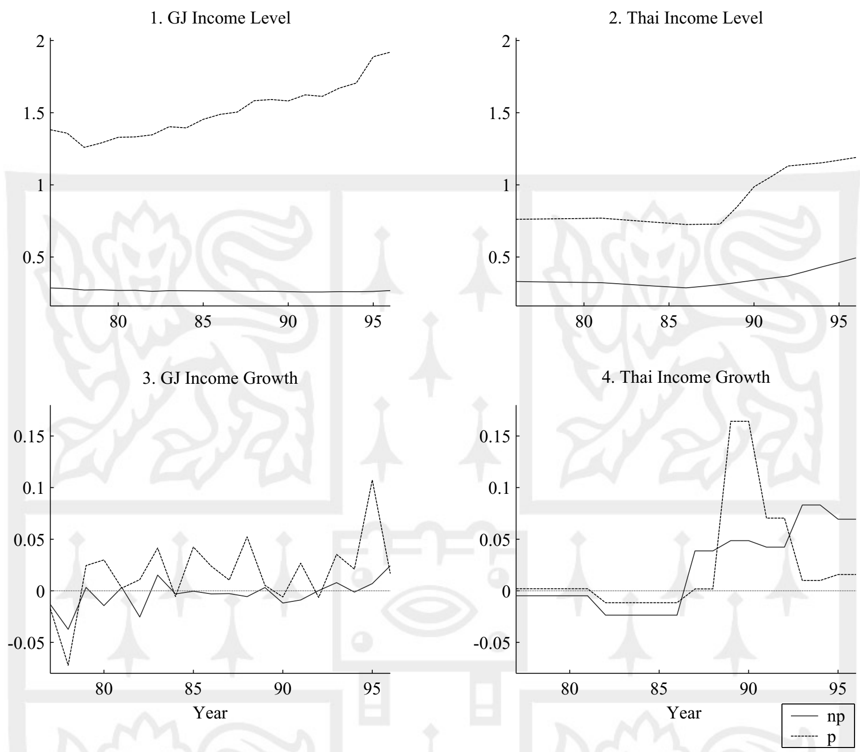

FIGURE 10. GJ subgroup growth dynamics.

in the model, and this is more or less true in the data, except for the catchup of nonparticipants after 1992. The average income grows for the participants but not for the nonparticipants in the model, while average income grows for both groups in the data. The co-movement in growth rates across participants and non-participants seems weak in the model, while it is strong in the data except for the catchup growth period after 1992. In particular, the model does not capture the growth-peak of participants during 1988–1992 and the catchup of nonparticipants after 1992.

Figure 11 compares the subgroup inequality dynamics between GJ and Thailand. Participants in the financial sector are richer than the non-participants in both the model and the data, a source of across-group inequality. The income ratio of the participant group to the non-participant group widens over time from 5 to 7 in the model, while it increases only moderately from 2.3 to 2.4 (peaking at 3.1 in 1992) in the data. The increase in inequality from the divergence in income levels across the two groups is a mirror image of growth features in GJ. The benefits of better investment are available only to the participants in the financial sector and they have a higher income growth than the nonparticipants. Incomes

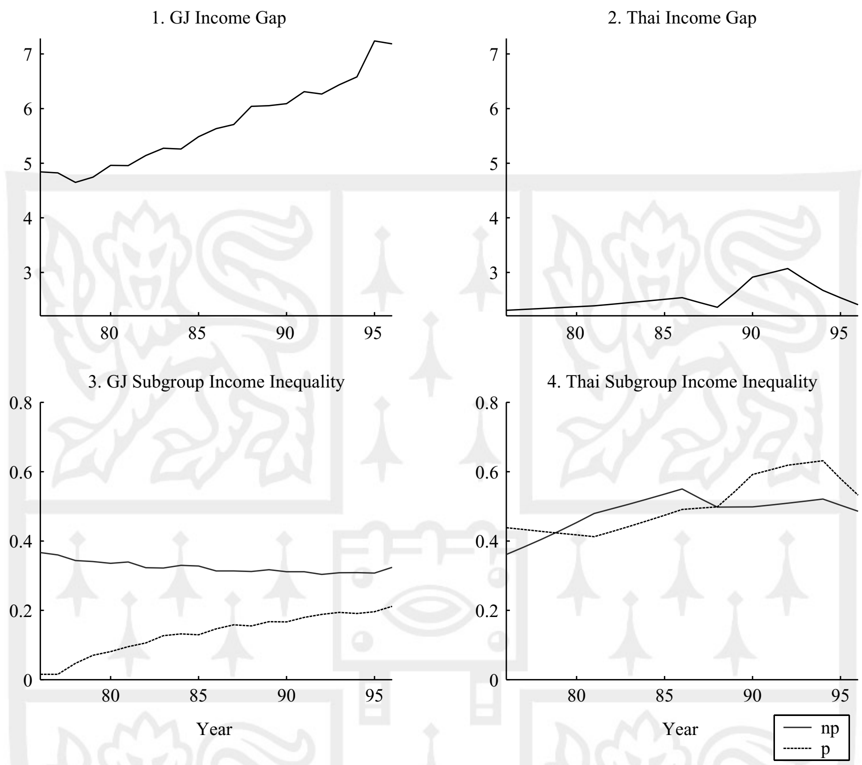

FIGURE 11. GJ subgroup inequality dynamics.

of nonparticipants also grow, but the rich households among them keep exiting to the participant group. These two effects within the nonparticipant group appear to offset one another in the model, keeping levels flat and average income growth among nonparticipants close to zero (see Figures 10.1 and 10.3).

GJ predicts that the poor group (nonparticipants) has higher inequality than the rich group (participants). However, inequality increases only among the participants, and it is stable (slightly decreasing) among the nonparticipants. These subgroup inequality features are related also to the entry-exit dynamics: the upper tail of the income distribution among nonparticipants is trimmed by exit, and the new entrants to the financial sector, poorer than the incumbents in the financial sector, are continually added to the lower tail of the income distribution of the participants. In contrast, in the data, there is no clear inequality ordering between the two groups and inequality increases for both of them, following a more or less common trend. This suggests that GJ seems to miss some driving forces of increase in inequality, common to both participants and nonparticipants, for example, educational expansion.[22]

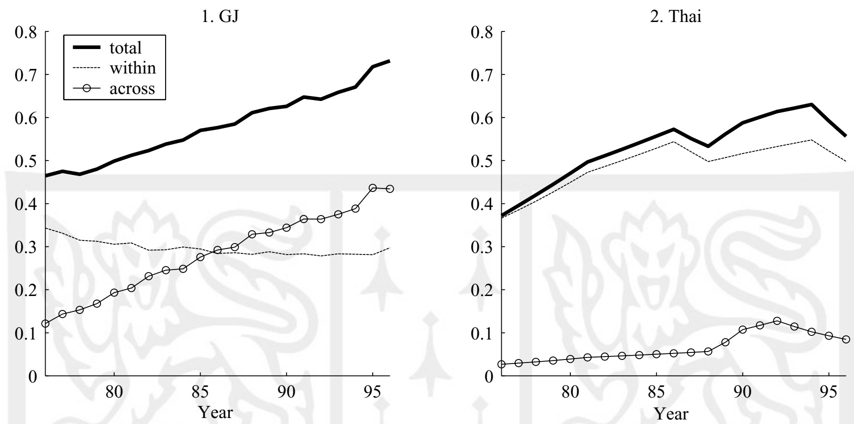

FIGURE 12. Within vs. across inequality decomposition.

The level of inequality is close to the data for the nonparticipant group, but it is much lower than the data for participant group. The diversification of idiosyncratic shocks in the financial sector of GJ seems excessive relative to the actual income variation among the participants in the data.

4.3.3. Decomposition. We apply the same decomposition methods in Section 2, with respect to financial sector participation. Total inequality is decomposed into within-group inequality and across-group inequality in Figure 12, comparing GJ to Thailand. In terms of both trends and patterns of movement over time, the driving force of inequality dynamics in GJ is across-group inequality, whereas it is within-group inequality in Thailand.

The two-decade growth rates of mean income and income inequality are decomposed in Tables 5 and 6, respectively. The model predicts a growth rate of mean income at 0.838, quite close to that of Thailand at 0.899. The composition effect of increasing participation in the financial sector on growth in GJ is substantial (at the rate of 0.654), as in the data (at the rate of 0.319). This composition effect is the main source of growth in GJ. Still, contribution of subgroup growth is larger than that of compositional growth in Thailand.

The model predicts an increase in income inequality over two decades, as in the data, but, with the predicted continual increase in inequality, the overall increase

TABLE 5. Decomposition of aggregate income growth in GJ

	Subgroup	Composition	Total
Thailand	0.580	0.319	0.899
GJ	0.184	0.654	0.838

TABLE 6. Decomposition of aggregate inequality change in GJ

	Within-Group		Across-Group		
	Subgroup	Composition	Income Gap	Composition	Total
Thailand	0.304	0.032	0.015	0.133	0.483
GJ	−0.014	−0.085	0.295	0.338	0.575

is higher in the model at 0.575 than in the data at 0.483. The composition effect, that is, the population shift from the nonparticipant group to the participant group, on across-group inequality change is substantial in the model (at the rate of 0.338) as in the data (at the rate of 0.133). As just noticed in Table 5, this population shift is also the source of income growth, so that the financial sector expansion is a significant link between growth and inequality dynamics both in the model and in the data. However, the effects in the model are exaggerated for both growth and inequality change; the orders of magnitudes in the model are larger than in the data. The divergence in income levels across the non-participant group and the participant group contributes to an increase in across-group inequality in the model at 0.295. The divergence is also observed in the data, but again with a much smaller order of magnitude at 0.015.

The composition effect contributes to a decrease in within-group inequality at low rate of −0.085 in the model because of population shifts from the high-inequality group (nonparticipants) to the low-inequality group (participants). The inequality ordering over the two groups is the opposite in the data, and so population shift contributes to an increase in within-group inequality. However, these effects are small in both model (−0.085) and data (0.032). The principal difference in inequality dynamics between model and data lies in the effect of changes in subgroup inequality on total inequality change. This effect is very small in the model but the most important source of inequality change in the data.

In sum, the effects of population shifts and income gaps across key groups in the model are overemphasized and subgroup effects are underemphasized relative to the data for both growth and inequality change. However, endogenizing financial participation in GJ solves the previous puzzles in LEB (where the participation was exogenously imposed), namely, lower income of the participants than the nonparticipants for entrepreneurs, convergence in income levels between the two groups, and a decrease in inequality among participants.

4.3.4. End-of-sample-period income distribution. Figure 13 compares the cumulative distribution function of income at the end of sample period, 1996, of the GJ model to that of the Thai data. The lower tail of the distribution is shifted to the left in the model as compared to the data. That is, the model predicts a high fraction of poor people relative to the data. The middle range of the distribution is flat in the model, that is, the model predicts a sparse middle class relative to

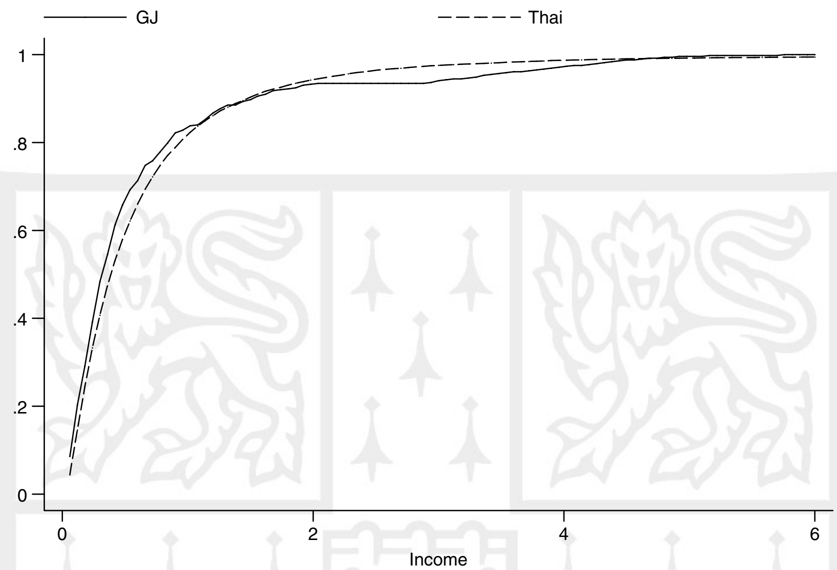

FIGURE 13. CDF of 1996 income distribution.

the data. This is due to the common fixed entry cost to the financial sector, that is, a common fixed sum of wealth is subtracted for new entrants into the financial sector. The distribution reaches unity at lower level of income in the model than in the data. That is, the model again does not capture the extremely rich people in the data.

We apply the Kolmogorov-Smirnov goodness-of-fit test to the GJ end-of-sample-period income distribution. The KS statistic is 0.58 (corresponding p-value is 0.89) and the null hypothesis of similarity of income distributions between the GJ model and the Thai data is accepted. Thus, despite the observed discrepancies above, overall shape of income distribution is quite well captured by the GJ model.

4.4. Sensitivity Analysis

All GJ parameters are estimated. To perform a sensitivity analysis, we vary the parameter values within one-standard-error bands around the point estimates. We fix γ at one because it is always pushed to this boundary value in estimation. Here bear in mind that for each variation of parameter values, the series of best-fitting realizations of aggregate shocks differs from the benchmark ones.

All features of GJ dynamics (income growth, inequality change, and financial expansion) are robust to the perturbation of preference parameter σ, fixed entry cost q, and idiosyncratic shock parameter $\overline{\epsilon}$ within their one-standard-error bands. An increase in σ makes the agents more risk averse and reduces the share of risky investment (that has a higher expected return) of the nonparticipants. This increases the value of risk-sharing as well as the income gap between participants and nonparticipants. Thus average income growth becomes smaller and inequality increase becomes larger. The effects of varying q are not certain *a priori*. As q increases, k^* also increases because the ratio of the two remains the same. This makes the economy wealthier, because of an increase in GJ scale in equation (40). An increase in q *per se* hinders non-participants from entering the financial sector, but the associated increase in wealth helps participation. Also an increase in q implies more lump sum payment of resources that are not used for investment. This may lower overall growth but the income gap between participants and nonparticipants may become smaller. Within the one-standard-error increase in q, the income growth rate is lower, inequality increases less, and participation rate is higher.

An increase in $\overline{\epsilon}$ implies an increase in the variance of idiosyncratic shock and makes the value of risk-sharing larger. In turn, participation in the financial sector is more attractive, but nonparticipants' income growth becomes lower. This results in lower average income growth and a higher rate of increase in inequality. However, the orders of magnitude of the changes in aggregate dynamics from perturbing parameters σ, q and $\overline{\epsilon}$ are small. Figure A.9 shows that by increasing $\overline{\epsilon}$ from 0.9954 to 1.0309, the annual average income growth rate decreases from 3.09% to 3.06%, and the rate of change in inequality increases from 2.30% to 2.72% per year. The end-of-sample-period fraction of participants increases from 23% to 27%.

Income growth and financial expansion turn out to be sensitive to aggregate shock parameters $\underline{\zeta}$, $\overline{\zeta}$ (which determine the rate of return to the risky investment), safe return δ, and discount factor β (which determine the utility return to saving), but inequality dynamics remain relatively robust. An increase in $\underline{\zeta}$ increases the expected return to risky investment and reduces the variance of aggregate shock. This makes the risky investment more attractive, and aggregate income growth and the shift to financial sector become faster. But this benefits participants more than nonparticipants, and income gap between them diverges faster. Figure A.10 displays that an increase in $\underline{\zeta}$ from 1.0470 to 1.0921 (the most sensitive case of all experiments) increases income growth rate from 3.09% to 5.99%, and makes income inequality grow at the rate of 4.69% per year. The end-of-sample-period fraction of participants increases to 36%.

An increase in $\overline{\zeta}$ increases the expected return to risky investment and its variance as well. Thus whether it makes the risky investment more attractive for the nonparticipants is uncertain, but it increases the value of advanced information on aggregate shock and hence the benefits from participation. Within the one-standard-error band, an increase in $\overline{\zeta}$ induces more income growth, faster financial expansion, and faster increase in inequality. An increase in δ benefits both participants and nonparticipants. Income growth is higher but the inequality dynamics remain about the same. The financial expansion is faster because of higher wealth accumulation on the part of nonparticipants. An increase in β makes the all agents more patient, and total investment increases. Its effects are similar to the increase in δ.

In summary, income growth and financial expansion are sensitive to parameters that determine the rates of return to investment, in particular the risky one, and less sensitive to the parameters of risk aversion, entry cost, and the idiosyncratic shock. The Ak nature of GJ technology seems to lie behind this result. Furthermore, the benefits of intermediation in the GJ model are driven more by the value of advanced information on aggregate shock (determined by $\underline{\zeta}$, $\overline{\zeta}$, and δ) than by the risk-sharing on the idiosyncratic shocks (determined by $\overline{\epsilon}$). Inequality dynamics remain robust to most perturbations.[23]

5. CONCLUSION

We evaluated two well-known macro models of growth and inequality which are built on explicit micro underpinnings and impediments to trade, that is, wealth-constrained self-selection. A two-step evaluation scheme was adopted. In the first step, we estimated most key parameters by fitting the assumed individual selection decisions of households without using the implied aggregate dynamic data. In the second step, we simulated the aggregate dynamics of growth and inequality together with subgroup dynamics of the models at those micro-fitted parameters and compared the predictions to the actual data. Thus, we use theories both for estimation and simulation. Explicit use of structural models and the framework of computational experiments, described by Kydland and Prescott (1996), helped us to organize our theoretical perspectives as well as the empirical data. However, explicit use of estimation also helped us to make consistent use of postulated economic environments and data. The importance of the latter has been emphasized by Hansen and Heckman (1996).

Not all available data were used in estimating parameters. Only the cross-sectional micro data on self-selection and wealth were used with likelihood methods to estimate key parameters. The aggregate dynamics and income distribution data were saved for testing. This separated use of data, estimation versus testing, helped us to avoid the potential danger of *ad hoc* overfitting, as addressed by Granger (1999). Our two-step micro-macro empirical strategy contributes to the synthesis between micro evidence and macroeconomic theory, envisioned by Browning, Hansen, and Heckman (1999).

Experimenting with two different models of growth and inequality and comparing them was helpful in identifying the salient patterns of the data and in documenting anomalies of the models. The parameter values chosen from cross-sectional micro estimation were not picked to generate nice aggregate dynamics. Surprisingly, however, the simulated aggregate dynamics of growth and inequality are close to the actual data for both models. In particular, LEB captures the aggregate movements of growth and inequality through endogenous factor price movements without aggregate shocks. GJ also does well with the long-run trend of growth with increasing inequality but not with the nonlinear patterns of income growth and inequality change.

Each model can predict compositional changes in the population across key selection groups as in the data. The effects of compositional changes on both income growth and income inequality change are substantial in each model and also in the data. This confirms Kuznets's (1955) hypothesis on the existence of a macro relationship between growth and inequality, but here *via micro channels* of wealth-constrained self-selection.

However, we also observed several anomalies. First, income gaps across key subgroups are too high in the models relative to the data. Second, neither model can predict the co-movement patterns across subgroups observed in the data. Third, neither model can replicate the movements of aggregate income growth and income inequality in relation to the growth patterns of entrepreneurs in the financial sector (the smallest but richest group). Fourth, income variation at the tails of the income distribution, in particular at the lower tail, are not well captured. The reasons for the anomalies are different across the two models, but the fundamental sources seem to be lack of appropriate aggregate shocks and heterogeneity. Still, from the comparative evaluation of two models, we learned that the simple addition of aggregate shocks and the introduction of more kinds of heterogeneity improve neither aggregate dynamics nor cross-sectional income distribution patterns. What matters is exactly *how* aggregate shocks and heterogeneity are incorporated, not the adding aggregate shocks or more kinds of heterogeneity itself. The GJ model, with aggregate shocks incorporated, captured the movements of aggregate income level and inequality worse than the LEB model that has no aggregate shocks. However, though having less kinds of heterogeneity, the GJ model with its endogenously structured costly access to the financial sector could remedy many of the anomalies of LEB.

Low within-group inequality among participants in financial sector suggests that insurance for idiosyncratic shocks in the model is overdone. Lack of co-movement across subgroup growth rates suggests that the informational advantage of financial sector in processing aggregate risk is too perfect. Ironically, high entrepreneurial income inequality among financial participants in LEB pushes one in the same direction. Thus, less perfect and less uniform financial markets for those with access are another promising feature to be incorporated.

We also learned that model specification and model evaluation are intimately related to each other. There is a thin edge between calibration and estimation that must be faced even if it appears in only one parameter, in our exercise in LEB. The support of idiosyncratic shock, the random setup cost, is bounded and fixed, as the shock enters the model in an additive way. Thus the choice of scale that converts wealth in the data into wealth in model units becomes important in both estimation and simulation. In fact, for some range of scales, there exists a trade-off between likelihood values from cross-sectional estimation and goodness-of-fit in simulated dynamics. However, once a scale parameter was chosen, most of key parameters of preferences and technology could be identified from explicit estimation with cross-sectional data alone. Thus, we were able to implement our two-step strategy.

We hope that our work here will enhance the synthesis between micro evidence and macroeconomic theory and allow an improved understanding of the relationship between growth and inequality. We also hope that our proposed model evaluation strategy can help to guide future research to advance the mutual penetration of quantitative economic theory and statistical observation.

NOTES

1. For a critical review of the diverse cross-country analysis results, see Jeong (2000) and Banerjee and Duflo (2000).

2. In fact, Jeong and Townsend (2007) choose the parameters of one of the benchmark models of this paper [a modified Lloyd-Ellis and Bernhardt (2000) model] matching the *aggregate dynamics* of output growth and factor shares (not the household choices) to explain the *components* of total factor productivity (TFP).

3. The regional average Gini coefficients are from Deininger and Squire (1996).

4. Banerjee and Newman (1993) and Aghion and Bolton (1997) also provide equilibrium theories of growth and inequality in relation to occupational choice and financial market imperfection, but in aggregate contexts. Jeong and Kim (2006) study complementarity between sector-specific experience and labor as another channel of earnings growth and inequality during transition to modern economic growth. They bring the model to the same Thai data applying a similar empirical strategy of simulation of aggregate dynamics of growth and inequality at the micro-fit parameters estimated from a cross-sectional earnings equation.

5. Specific form of z function is derived in Appendix A.2.

6. This limited use of data helps us to avoid the following endogeneity issue. Suppose we observe a wealthy household whose occupational choice is an entrepreneur. Then, it can be either the case that the household is wealthy because of its previous occupational choice as entrepreneur, or the case that it chose to be an entrepreneur because of its high initial wealth. Using the entire sample, we cannot distinguish between the two cases.

7. To check the robustness of the estimates, we varied the sample of young households by changing the cutoff age to 25 and also to 35 and got similar estimates. As the average age of Thai household heads is 45, this range of variation in identifying young households seems reasonable.

8. The bootstrap was run with 10,000 random resamplings with replacement. The standard errors are reported in scale of 10^{-12}.

9. A referee suggests that the relative precision levels across the estimates appear to be counterintuitive, observing that the curvature parameters ξ, ρ and σ are much more precisely estimated than the slope parameters α and β. Although the referee mistakenly categorize between curvature and slope parameters (in fact, β, ρ and σ are curvature parameters and α and ξ are slope parameters), this is an interesting question worth to discuss, leading to a better understanding of the nature of the LEB estimation. Related, the referee also seeks why we have $\alpha \simeq 1$ and $\sigma = 0$. Note that the likelihood is written on occupational choices, not directly on production function. Thus, it is the three curvature parameters $x^*(w)$, $b^*(w)$ and $\widetilde{b}(w)$ of the critical setup cost function $z(b, w)$ (or equivalently the curvature parameters of profit function, see Appendix A.2) in relation to the occupational choice data, not those of production function, which determine the precision as well as the values of the estimates. The parameters ξ and ρ are precisely estimated because $\widetilde{b}(w) = \frac{(\xi - w)^2}{2\rho} - w$ is precisely estimated. α is estimated close to one σ is (tightly) estimated at zero because the gap $x^*(w) - \widetilde{b}(w) = \frac{\rho(\alpha-1)}{2(\rho\beta - \sigma^2)} + \sigma \frac{\rho(\xi - w)}{2(\rho\beta - \sigma^2)}$ is estimated very small and does not depend on wage w, implying $\alpha \simeq 1$ and $\sigma = 0$. These technology parameters estimates (in particular $\sigma = 0$) obtained fitting the micro occupational choices put restrictions on aggregate dynamics.

10. This can happen as a result of the myopic nature of the LEB preferences.

11. We normalize the Thai income unit by matching the 1976 mean income levels between the model and the data for the sake of convenient comparison. This normalization does not affect the inequality levels in the data.

12. The decomposition of ΔAI involves a log approximation. See Mookherjee and Shorrocks (1982).

13. The discrepancy between the sum of component changes and the total change is a result of the log approximation in (12).

14. We generate the empirical distribution functions with sample size of 100 for both the LEB model and the Thai data. The p-value is from Smirnov (1948).

15. In fact, the dynamics are robust to the variation of α, β, and σ over the *entire* ranges, not just within 10-percent-deviation ranges, of the parameter space that satisfy the restrictions of the LEB model, which are $[1, 1.3]$ for α, $[0.01, \infty)$ for β, and $[0, 0.013]$ for σ.

16. See Appendix A.2.

17. In their original model, GJ consider a log utility function, a special case of the CRRA preferences with $\sigma \to 1$.

18. See Townsend and Ueda (2006) for full discussion.

19. The latter two lines in (39) define zero-probability events of corner solutions.

20. The $\widehat{k}$, the top 6.4 percentile wealth in 1976, is 204,824 baht in the data.

21. Again, we normalize the Thai income unit by matching the 1976 mean income levels between the model and the data.

22. See Jeong (2000) for the important role of educational expansion on increase in inequality in Thailand.

23. In contrast, in LEB, inequality dynamics were much more sensitive than income growth and occupation choice dynamics.

24. In the 1976 SES, financial income was not recorded separately and thus cannot be added to the total income for this year.

25. There are other asset items recorded in the SES, depending on years. We choose the items that are commonly collected across all sampling years.

26. This single proxy variable accounts for 26 to 36% of the total variation of the ownership of sixteen assets. The scoring coefficients of the first principal component are available upon request.

27. To calculate the representative interest rate, we use the time-series of average of the lending rates of commercial banks, finance companies, and interbank loans between 1978 and 1996. (Data source: Economic Research Department at the Bank of Thailand.)

REFERENCES

Aghion, Philippe and Patrick Bolton (1997) A theory of trickle-down growth and development. *Review of Economic Studies* 5(64), 151–172.

Banerjee, Abhijit and Esther Duflo (2000) Inequality and Growth: What Can the Data Say? NBER Working Paper 7793.

Banerjee, Abhijit and Andrew F. Newman (1993) Occupational choice and the process of development. *Journal of Political Economy* 5(101), 274–298.

Bourguignon, François (2002) The Distributional Effects of Growth: Case Studies vs. Cross-country Regression. DELTA Working Paper 2002–23.

Browning, Martin, Lars P. Hansen, and James J. Heckman (1999) Micro data and general equilibrium models. In John B. Taylor and Michael Woodford (eds.), *Handbook of macroeconomics*, Vol. 1A. Amsterdam: North Holland.

Deininger, K. and L. Squire (1996) A new data set measuring income inequality. *World Bank Economic Review* 5(10), 565–591.

Frisch, Ragnar (1933) Editorial. *Econometrica* 5(1), 1–4.

Gine, Xavier and Robert M. Townsend (2004) Evaluation of financial liberalization: A general equilibrium model with constrained occupational choice. *Journal of Development Economics* 74, 269–307.

Granger, Clive W.J. (1999) *Empirical modeling in economics: Specification and evaluation*. Cambridge: Cambridge University Press.

Greenwood, Jeremy and Boyan Jovanovic (1990) Financial development, growth, and the distribution of income. *Journal of Political Economy* 5(98), 1076–1107.

Hansen, Lars P. and James J. Heckman (1996) The empirical foundations of calibration. *Journal of Economic Perspectives* 5(10), 87–104.
Jeong, Hyeok (2000) *Sources of Kuznets Dynamics in Thailand*. Ph.D. dissertation. University of Chicago.
Jeong, Hyeok (in press) Assessment of relationship between growth and inequality: Micro evidence from thailand. *Macroeconomic Dynamics*.
Jeong, Hyeok and Yong Kim (2006) Complementarity and Transition to Modern Economic Growth. IEPR Working Paper Series 06.44.
Jeong, Hyeok and Robert M. Townsend (2007) Sources of TFP growth: Occupational choice and financial deepening. *Economic Theory* 32(1), 179–221.
Kuznets, Simon (1955) Economic growth and income inequality. *American Economic Review, Papers and Proceedings* 5(45), 1–28.
Kydland, F.E. and Edward. C. Prescott (1982) Time to build and aggregate fluctuations. *Econometrica* 5(50), 1345–1370.
Kydland, F.E. and Edward. C. Prescott (1996) The computational experiment: An econometric tool. *Journal of Economic Perspectives* 5(10), 69–85.
Lloyd-Ellis, Hew and Dan Bernhardt (2000) Enterprise, inequality, and economic development. *Review of Economic Studies* 5(67), 147–168.
Mookherjee, Dilip and Anthony F. Shorrocks (1982) A decomposition analysis of the trend in UK income inequality. *Economic Journal* 5(92), 886–902.
Oreskes, Naomi, Kristin Shrader-Frechette, and Kenneth Belitz (1994) Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 5(263), 641–646.
Smirnov, N. (1948) Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics* 5(19/2), 279–281.
Townsend, Robert M. (1978) Intermediation with costly bilateral exchange. *Review of Economic Studies* 5(45), 417–425.
Townsend, Robert M. and Kenichi Ueda (2006) Transitional growth with increasing inequality and financial deepening. *Review of Economic Studies* 73(1), 251–293.

APPENDIX

A.1. DATA DESCRIPTION

We use the Thai Socio-Economic Survey (SES), a nationally representative micro survey conducted by the National Statistical Office in Thailand. Over the recent two decades, between 1976 and 1996, eight rounds of repeated cross-sections were collected: 1976, 1981, 1986, 1988, 1990, 1992, 1994, and 1996, using a clustered random sampling, stratified by geographic regions over the entire country. The sampling unit is household and the sample size varies from 10,897 to 25,208, becoming larger in later years. Economically active households in the SES data were selected for our analysis.

The income is measured in real annual terms in 1990 baht value, which includes earned income (profit income for the self-employed and wage income for the workers) and financial income (land rents, interest and dividend income from financial investment).[24]

For a binary occupation category, we call entrepreneurs by the nonfarm business households and the rest are categorized into nonentrepreneurs. For a financial participation category, we identify the participants with the households who actually made financial transactions with any of the formal financial intermediaries—commercial banks, a government savings banks, the Bank of Agriculture & Agricultural Cooperatives (BAAC), a government housing bank, financial companies, or credit financiers, and the rest are nonparticipants.

The SES does not directly record total wealth data, but it does record various wealth items. Using the information on these wealth items, we construct a proxy for household wealth. We first use the information on ownership of sixteen household assets: private water supply, gasoline-cooking equipment, access to electricity, phone, sofa, bed, stove, refrigerator, electric iron, electric pot, radio, TV, motorcycle, car, sewing machine, and motor boat.[25] Applying principal-component analysis to these variables, we pick the first-principal component, which best summarizes the variations of the ownership of the sixteen assets by a single variable, as a proxy for the latent wealth underlying them.[26] This asset index is not in monetary units. To convert it into monetary unit, we use the rental value of owned house as the representative rental price and multiply the above asset index by the rental price. This gives an approximate estimate of *flow value* of the above sixteen assets. We add other rental incomes to this to get the total flow value of wealth. We then divide this flow value of wealth by the interest rate to get an estimate of *stock value* of wealth.[27]

A.2. LEB IDENTIFICATION

For the quadratic form of technology,

$$f(k,l) = \alpha k - \frac{\beta}{2}k^2 + \xi l - \frac{\rho}{2}l^2 + \sigma lk,$$

labor demand l is linear in capital demand k

$$l = \frac{\xi - w}{\rho} + \frac{\sigma}{\rho}k, \qquad \textbf{(A.1)}$$

and profit function can be expressed as a second-degree polynomial of capital demand k:

$$\pi(b,x,w) = C_0(w) + C_1(w)k + C_2k^2 - x, \qquad \textbf{(A.2)}$$

where

$$C_0(w) = \frac{(\xi - w)^2}{2\rho}, \qquad \textbf{(A.3)}$$

$$C_1(w) = \alpha - 1 + \frac{\sigma(\xi - w)}{\rho}, \qquad \textbf{(A.4)}$$

$$C_2 = \frac{1}{2}\left(\frac{\sigma^2}{\rho} - \beta\right). \qquad \textbf{(A.5)}$$

Capital demand k depends on wealth b as well as technology parameters $(\alpha, \beta, \xi, \rho, \sigma)$ and market wage w, if the entrepreneurs are constrained. Unconstrained capital demand k^*, independent from wealth, is given by

$$k^* = \frac{C_1(w)}{-2C_2}.$$

Note that the key determinant in occupational choice is the critical setup cost function $z(b, w)$, which is derived by equating the above profit function in (A.2) with wage w. Let $b^*(w)$ be the critical level of wealth above which the wealth constraint does not bind in occupational choice, $x^*(w)$ be the associated level of critical setup cost, and $\widetilde{b}(w)$ be the wealth level below which the wealth constraint binds exactly at the level of setup cost (hence the capital demand hits the lower bound zero). These three objects $b^*(w)$, $x^*(w)$, and $\widetilde{b}(w)$, conditional on wage w, fully characterize the occupational choice in the LEB model and the z function is given as.

$$\begin{aligned} z(b, w) &= x^*(w), \quad \textit{if } b \geq b^*(w) \qquad \textbf{(A.6)} \\ &= b + \frac{C_1(w) + 1 - \sqrt{(C_1(w)+1)^2 - 4C_2\,(C_0(w) - b - w)}}{2C_2}, \\ &\qquad\qquad \text{if } \widetilde{b}(w) \leq b < b^*(w) \\ &= b, \quad \textit{if } b < \widetilde{b}(w), \end{aligned}$$

where

$$\widetilde{b}(w) = C_0(w) - w, \qquad \textbf{(A.7)}$$

$$x^*(w) = \widetilde{b}(w) - \frac{C_1(w)^2}{4C_2}, \qquad \textbf{(A.8)}$$

$$b^*(w) = x^*(w) - \frac{C_1(w)}{2C_2}. \qquad \textbf{(A.9)}$$

That is, the three coefficients $C_0(w)$, $C_1(w)$, and C_2, of profit function determine not only the income of entrepreneurs but also the occupational map. Thus, the LEB log likelihood function can be reduced into the following form:

$$\log L(\alpha, \beta, \xi, \rho, \sigma, m; w) = \log L(C_0, C_1, C_2, m; w)$$

and only three out of five production parameters can be identified if a *single* wage is used.

Unlike the models with a Cobb-Douglas production function, factor income shares are not stationary in LEB and they cannot be used in pinning down the remaining technology parameters. However, by *adding variation in wage*, we can uncover the full five production parameters as follows. Suppose we use wage variation over time, say between two initial years, 1976 and 1981. Given the wage in 1976, w^{76}, the critical setup cost function z and the log likelihood function are characterized by the three coefficients $C_0(w^{76})$, $C_1(w^{76})$, and C_2, and similarly the $C_0(w^{81})$, $C_1(w^{81})$, and C_2, at the wage in 1981, w^{81}. Note that the coefficient C_2 does not depend on wage, and should be the same over time. This plays a role of an identifying restriction. We form the log likelihood function over the two years

$$\log L\left(C_0^{76}, C_1^{76}, C_2, m; w^{76}\right) + \log L\left(C_0^{81}, C_1^{81}, C_2, m; w^{81}\right), \qquad \textbf{(A.10)}$$

where $C_0^{76} = C_0(w^{76})$, $C_1^{76} = C_1(w^{76})$, $C_0^{81} = C_0(w^{81})$, and $C_1^{81} = C_1(w^{81})$. By allowing exogenous growth in the subsistence income γ, the reservation wage level will also exogenously vary over time. Thus we can avoid the endogeneity problem in our estimation using two different wage rates at initial years, during which the slow wage growth in Thailand is considered as the exogenous reservation wage growth.

Now we estimate C_0^{76}, C_1^{76}, C_0^{81}, C_1^{81}, and C_2 by maximizing the log likelihood function in (A.10). Then, the five production parameters $(\alpha, \beta, \xi, \rho, \sigma)$ can be identified as follows. First, from dividing C_0^{76} by C_0^{81}, we find ξ:

$$\xi = \frac{w^{81}\sqrt{C_0^{76}} - w^{76}\sqrt{C_0^{81}}}{\sqrt{C_0^{76}} - \sqrt{C_0^{81}}}. \tag{A.11}$$

Then, substituting this ξ either into C_0^{76} or into C_0^{81}, we find ρ:

$$\rho = \frac{1}{2}\left(\frac{w^{81} - w^{76}}{\sqrt{C_0^{76}} - \sqrt{C_0^{81}}}\right)^2. \tag{A.12}$$

Using this ρ and subtracting C_1^{76} from C_1^{81}, we get σ:

$$\sigma = \frac{1}{2}\frac{(w^{81} - w^{76})\left(C_1^{76} - C_1^{81}\right)}{\left(\sqrt{C_0^{76}} - \sqrt{C_0^{81}}\right)^2}. \tag{A.13}$$

Substituting these ξ, ρ, and σ either into C_1^{76} or into C_1^{81}, α can be found:

$$\alpha = 1 + \frac{C_1^{81}\sqrt{C_0^{76}} - C_1^{76}\sqrt{C_0^{81}}}{\sqrt{C_0^{76}} - \sqrt{C_0^{81}}}. \tag{A.14}$$

Finally, substituting ρ, and σ into C_2, we get β:

$$\beta = \frac{1}{2}\left(\frac{C_1^{76} - C_1^{81}}{\sqrt{C_0^{76}} - \sqrt{C_0^{81}}}\right)^2 - 2C_2. \tag{A.15}$$

The support of the setup cost x is specified as unit interval, and the critical values $\widetilde{b}$ and x^* in the z function should satisfy the following relations:

$$0 \le \widetilde{b}(w^t) \le 1,$$
$$0 \le x^*(w^t) \le 1,$$

which implies that C_0^t, C_1^t, and C_2 should satisfy

$$C_0^t - w^t \ge 0, \tag{A.16}$$

$$C_0^t - w^t - \frac{C_1^{t\,2}}{4C_2} \le 1. \tag{A.17}$$

Furthermore, since z is increasing and concave in b, the following restrictions should be met:

$$\widetilde{b}(w^t) \leq x^*(w^t),$$
$$x^*(w^t) \leq b^*(w^t),$$

which again implies that

$$C_2 \leq 0, \tag{A.18}$$

$$C_1^t \geq 0. \tag{A.19}$$

The inequality constraints from (A.16) to (A.19) restrict the LEB parameter space and are imposed in estimation.

A.3. LEB SIMULATION ALGORITHM

The LEB model is simulated using the algorithm of Gine and Townsend (2004). At every date there is a distribution of beginning-of-period wealth, presumed to lie on some a priori grid. Guessing a wage, along with the parameters of technology, the regions of the occupation partition are pinned down. The distribution of talent then determines the fractions of the population choosing to be workers, subsisters, or entrepreneurs at each level of wealth. Adding up over all wealth levels, these population fractions should sum to one, and otherwise the labor market does not clear. This procedure is repeated to find an equilibrium wage in a bisection algorithm. Thus, end-of-period wealth is determined. A fraction ϖ of this wealth is saved, and this determines next period's distribution of beginning-of-period wealth. The distribution of setup cost for entrepreneurs adds additional diversity. The lower end point of the wealth distribution is the wealth of the household in the previous period who had least beginning-of-period wealth and the lowest talent (highest setup cost), and the upper end point is associated with the household in the previous period who had the highest beginning-of-period wealth and the highest talent (lowest setup cost). The initial condition of the model is the estimated initial distribution of wealth. Here we take the 1976 SES wealth distribution, scaled by the chosen wealth scale used in the estimation, as the initial wealth distribution for simulation. One period in the simulation corresponds to one year in the data.

In the data, there are households who indeed have access to financial sector. Because the original LEB model does not distinguish between the participants and non-participants in the financial sector, it is modified to include an exogenously embedded intermediated sector. Those in the intermediated sector can borrow and lend their wealth at an equilibrium interest rate, determined in a bisection algorithm. In this sector, the occupational map of Figure 1 is completely flat. That is, wealth does not determine occupational choice or scale of enterprise. There is only a common critical value for setup cost. This sector is otherwise integrated with the rest of the economy via a common labor market and hence a common market wage. Thus, the wage and interest rate are determined simultaneously. At each period, the number of households in the intermediated sector is specified exogenously, and made to increase at the observed rate of increase in participation as in the SES data, from 6% in 1976 to 26% in 1996.

A.4. GJ SIMULATION ALGORITHM

The GJ model is simulated using the algorithm of Townsend and Ueda (2006). The burden here is finding the value functions v^0 and v^1 described earlier, which are not bounded, and the support of which is evolving over time. Still, the value function v^1 has a closed form solution up to the risk aversion parameter σ, as does a value function for a fictitious sector, those never allowed to enter the financial system. The value function v^0 can be trapped between these two. The non-convex aspect of the problem, the one associated with the fixed entry cost q, disappears in the limit, as the horizon is driven to infinity. The value functions converge after iteration. The space of value functions is reasonably well approximated by Chebyshev polynomials. Policy functions for investment and portfolio share are found by grid search with successive refinements.

The dynamic path of the GJ simulation depends on the realization of aggregate shocks. Thus we need to choose a specific path of realized outcomes of aggregate shocks to determine which GJ simulation is to be compared with the data. Here we pick the path that is closest, out of 500 Monte Carlo simulated paths at the chosen parameter values, to the Thai aggregate dynamics according to a root-mean-squared-error (RMSE) metric, defined over the aggregate paths of income growth rate, income inequality level, and fraction of participants, equally weighted.

A.5. PROFILES OF WEALTH AND AGGREGATE DYNAMICS

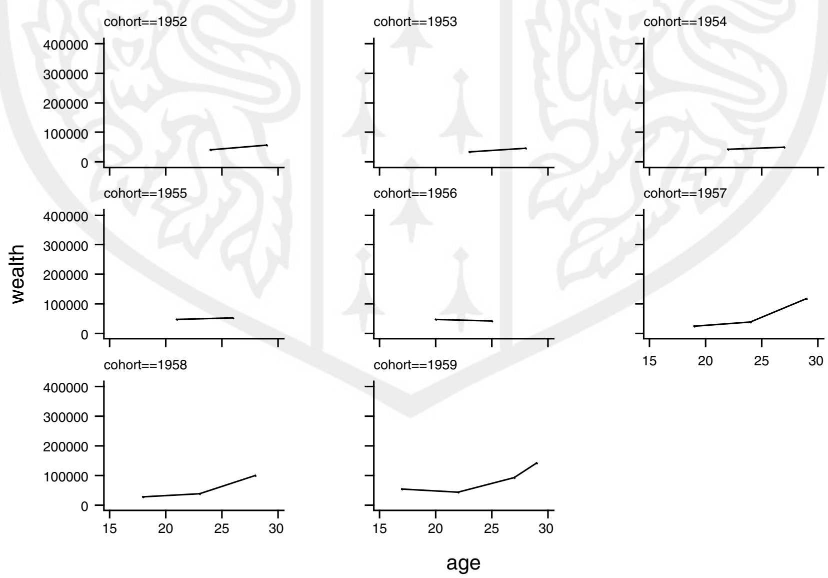

FIGURE A.1. Thai age profile of wealth by cohort: age < 30.

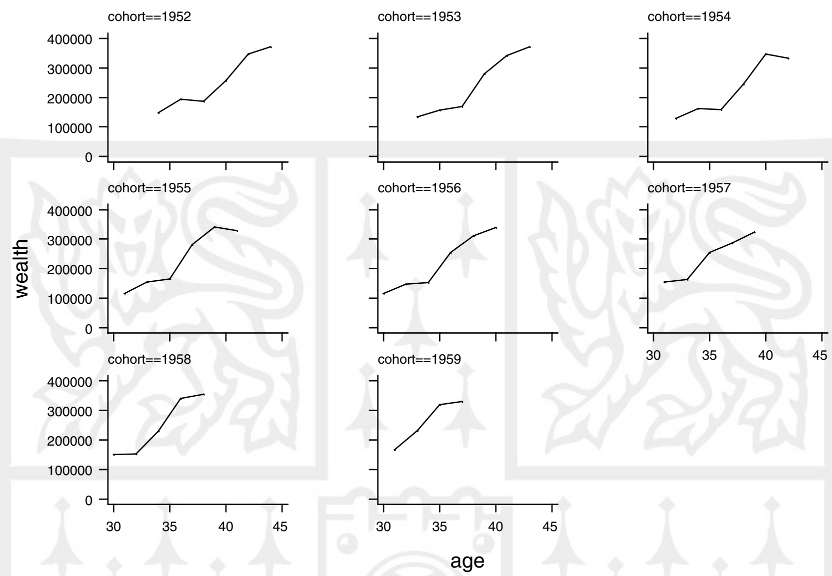

FIGURE A.2. Thai age profile of wealth by cohort: age > = 30.

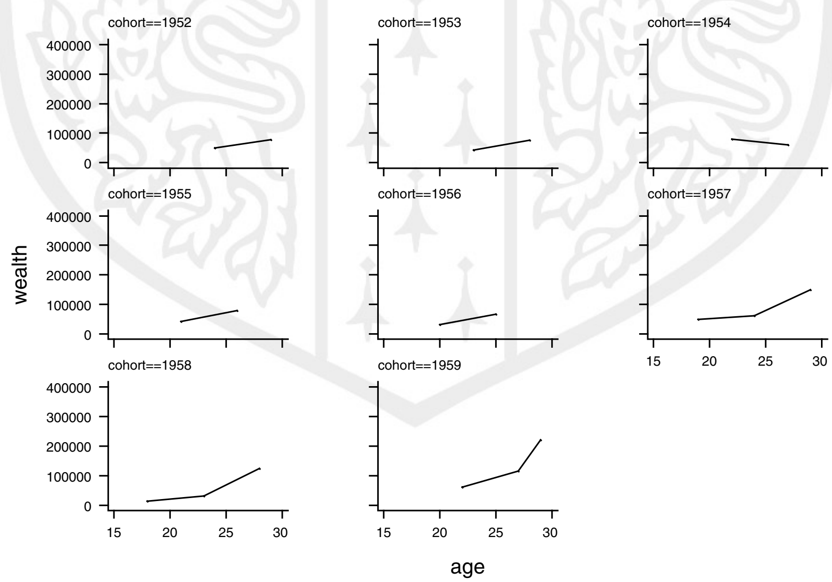

FIGURE A.3. Thai age profile of wealth of entrepreneurs by cohort: age < 30.

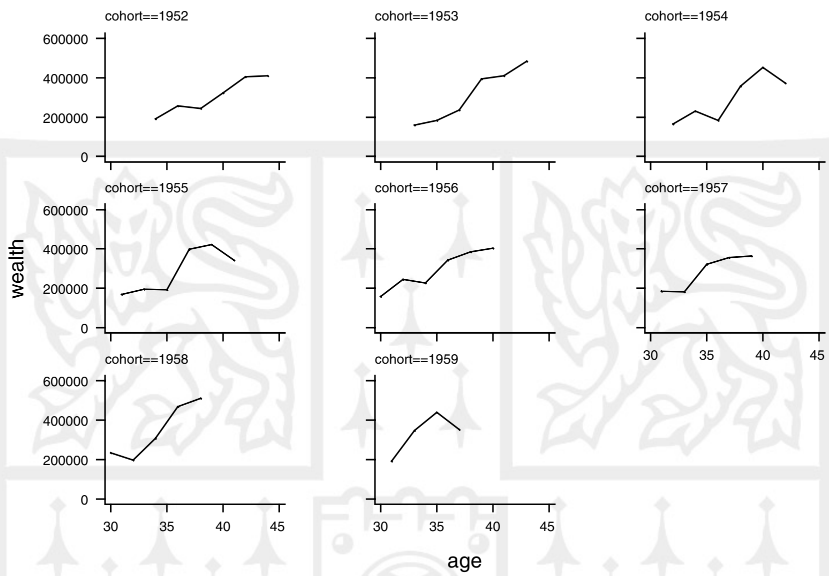

FIGURE A.4. Thai age profile of wealth of entrepreneurs by cohort: age $>=30$.

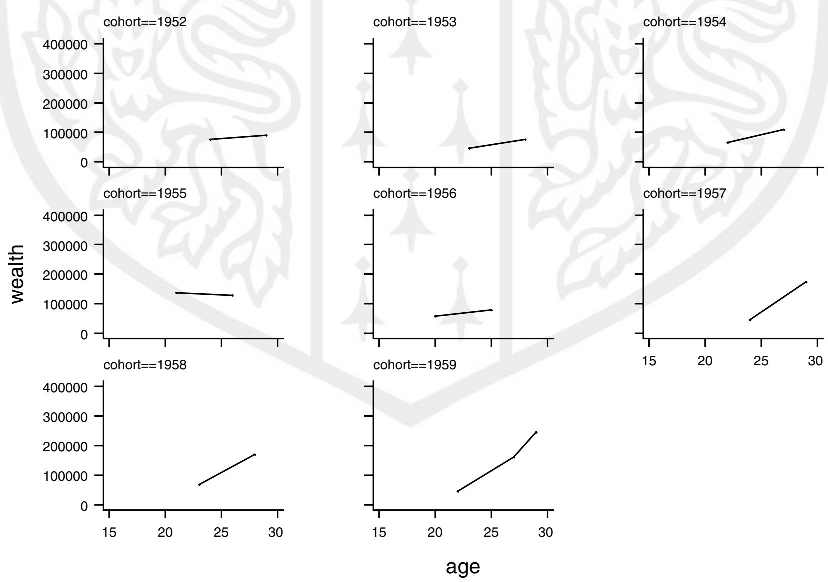

FIGURE A.5. Thai age profile of wealth of participants by cohort: age < 30.

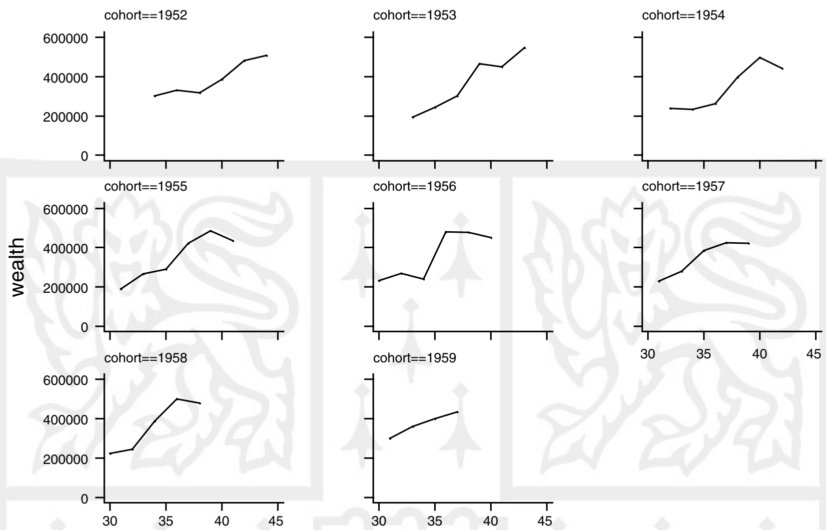

FIGURE A.6. Thai age profile of wealth of participants by cohort: age > = 30.

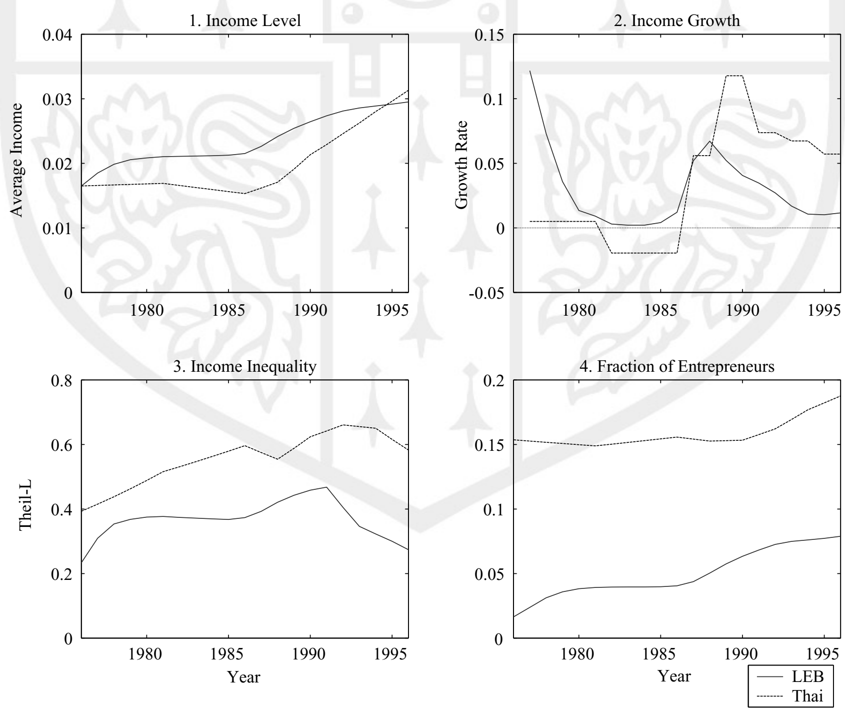

FIGURE A.7. LEB aggregate dynamics at alpha = 1.3.

xxx mdy07014 September 30, 2007 15:45

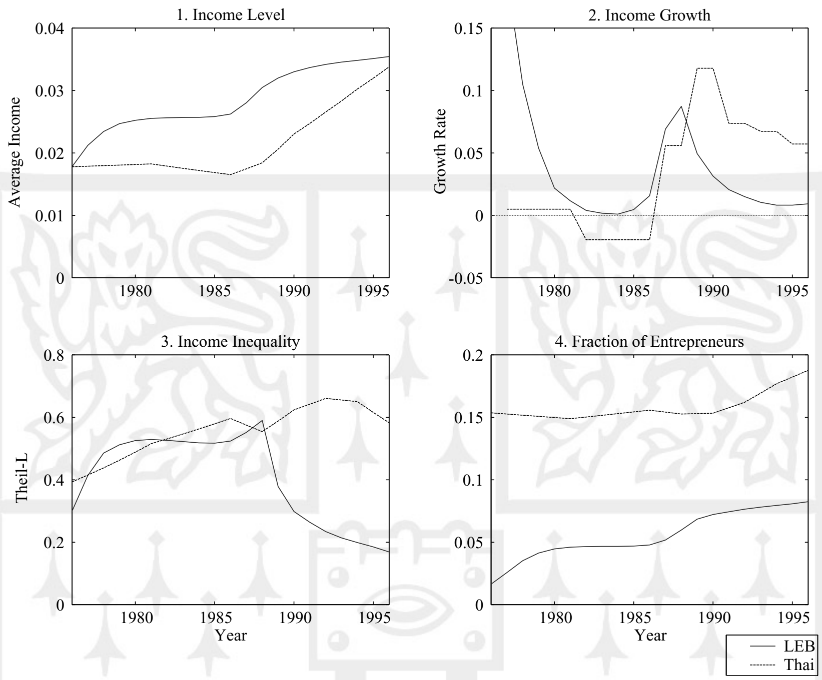

FIGURE A.8. LEB aggregate dynamics at xi = 0.0623.

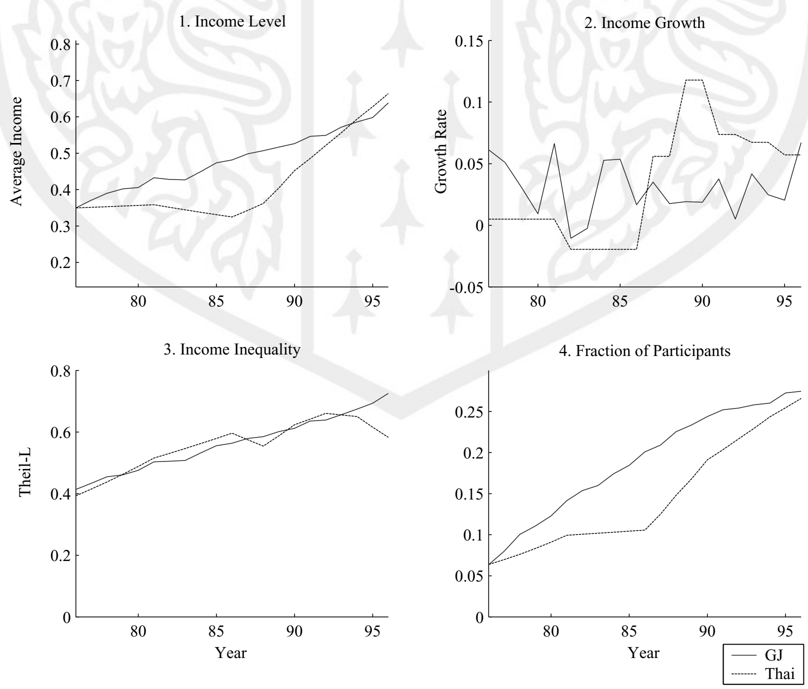

FIGURE A.9. GJ aggregate dynamics at idiosyncratic shock upper bound = 1.0309.

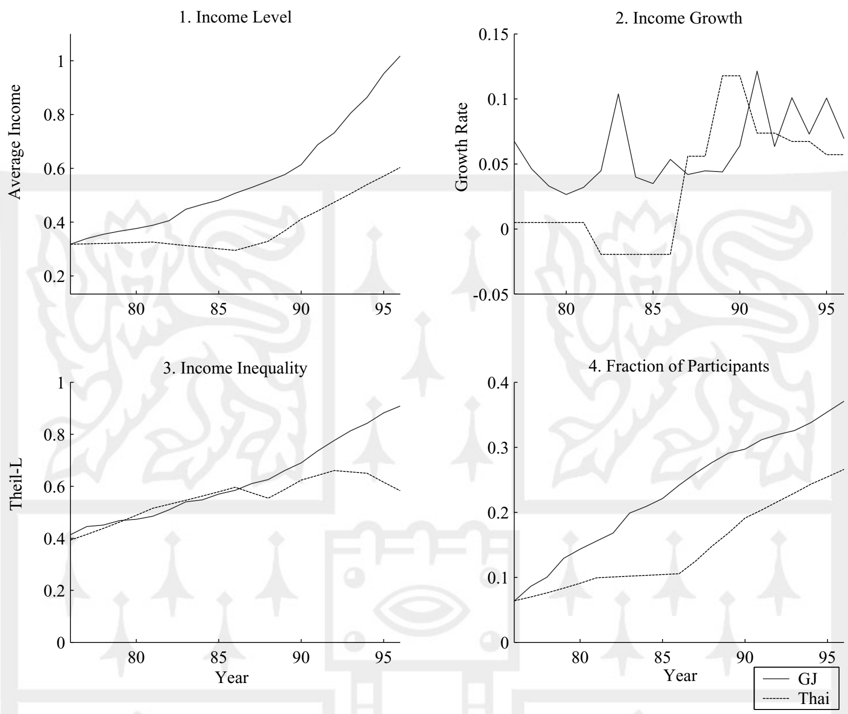

FIGURE A.10. GJ aggregate dynamics at aggregate shock lower bound = 1.0921.