Financial Frictions and the Persistence of History: A Quantitative Exploration

Francisco J. Buera* Yongseok Shin†

September 2007

Abstract

This paper quantifies the role of financial frictions in economic development. We incorporate financial frictions and capital misallocation into an otherwise-standard growth model, and calibrate by matching its stationary equilibrium to the data on standard macroeconomic aggregates, firm-size distribution and firms' external financing. We find that financial frictions have small effects on the output per capita in the stationary equilibrium because most entrepreneurial individuals accumulate wealth and overcome the frictions over time. However, financial frictions do have a large and persistent impact along the transition to the steady state, especially when capital is misallocated initially. In fact, financial frictions and capital misallocation are keys to understanding the observed economic transitions that are not explained by the standard neoclassical model. Our model economy converges slowly to the steady state, with the interest rate, investment rate and total factor productivity starting low and rising over time.

Keywords: Financial frictions; Development dynamics; Aggregation

*Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, USA (e-mail: f-buera@northwestern.edu).

†Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706, USA (e-mail: yshin@ssc.wisc.edu).

Explaining cross-country differences in economic development is a never-ending quest for economists. Recent quantitative research has pointed out some important pieces of this puzzle: Cross-country income differences are mainly accounted for by the low total factor productivity (TFP) in poor countries (Hall and Jones, 1999; Klenow and Rodríguez-Clare, 1997); The misallocation of resources across productive units is an important source of the low aggregate productivity in less-developed countries (Hsieh and Klenow, 2007; Restuccia and Rogerson, 2007); TFP growth and sectoral reallocation have been largely responsible for recent growth miracles (Hsieh, 2002; Young, 1995). As the fundamental role of credit markets is to deploy economic resources to their most productive use, it is natural to ask whether these development facts can be explained by cross-country differences in the degree of financial market imperfections.

Indeed, it has been widely recognized that well-functioning credit markets are an important institution missing from poor economies (Banerjee and Duflo, 2005), and a large theoretical literature has considered the role of credit markets in economic development.[1] This literature constructs tractable models to obtain equilibrium descriptions analytically, and concludes that with financial frictions the initial distribution of resources across individuals has large and permanent effects. In spite of such theoretical contributions, very few quantitatively-oriented attempts have been made to explore whether the micro-level distortions caused by financial frictions have significant aggregate effects. Do financial frictions still have large effects in the long run if agents can save to overcome these constraints? Do imperfect credit markets have a large impact on the dynamics of development? Our paper provides quantitative answers to these questions.

We incorporate financial frictions into an otherwise-standard growth model, and calibrate by matching its stationary equilibrium to the US data on standard macroeconomic aggregates, firm-size distribution, and firms' external financing, among others. Economies in our analysis will be different from one another in their degree of financial frictions. This approach allows us to assess the quantitative implications of financial frictions on economic development in an empirically-relevant setup.

In our model, individuals choose whether to operate an individual-specific technology—become entrepreneurs—or to supply labor for a wage in each period. Individuals differ in their productivity as an entrepreneur and in their wealth, with the latter being endogenously determined by forward-looking saving decisions to maximize their dynastic utility. We introduce financial frictions in the form of collateral constraints that inhibit efficient re-allocation of capital across entrepreneurs. In particular, a talented would-be entrepreneur who is born poor will have to work for a wage until she has accumulated enough net worth to overcome the financial frictions and operate her technology at an efficient scale. At any given point in time, hence, there is some inefficient use of resources, which we call misallocation.

Our first result pertains to the long run: Financial frictions alone have a small impact on the output per capita in the stationary equilibrium, belying the predictions of the theoretical literature. As we tighten the financial constraints, steady-state output drops by at most 10 percent

[1] See Banerjee and Duflo (2005) and Matsuyama (2007) for recent surveys on this literature.

relative to the benchmark economy calibrated to the US. This result hinges on the transitory nature of financial frictions: A talented but poor individual will accumulate enough wealth quickly so that she can overcome the financial constraints and operate her productive technology at the maximal-profit scale. The underlying intuition is similar to why the standard neoclassical growth model converges very quickly to the steady state: With high marginal returns to capital, the saving rate is correspondingly high. Therefore, in the resulting stationary equilibrium, only a small fraction of individuals are bound by the constraints. Self-financing, while potentially costly from an individual's perspective (Buera, 2006), is a very good substitute to formal financial markets in the context of the macroeconomy.[2]

Our second result shows that financial frictions do have a substantial and enduring impact on the transitional dynamics of the macroeconomy. In fact, financial frictions and capital misallocation are keys to understanding the observed economic transitions that are not explained by the standard neoclassical model. With misallocation of initial resources, an economy produces less and invests less than it would otherwise. Over time, this misallocation is unwound and the economy grows more productive. Financial frictions delay such efficient reallocation, and prolong the impact of the initial misallocation. One consequence is that the economy converges to the steady state at a slow pace. In our examples with misallocation of the initial capital, the half-life of the aggregate capital stock with financial frictions is typically two to three times that in the perfect-credit economy.

The evolution of the ability-wealth allocation over time generates endogenous dynamics for TFP. Unlike in the standard growth model, our economy's aggregate output is a function of the entire joint distribution of wealth and entrepreneurial talent, beyond the aggregate capital stock (unconditional first moment). An economy with severe misallocation of wealth and entrepreneurial talent will have a lower aggregate output than another with the same aggregate capital but a better allocation of wealth to talent. As the initial misallocation is unwound over time, the endogenous dynamics of the ability-wealth distribution in the model is reflected on the imputed TFP series. In our examples with initial misallocation, the model generates TFP growth of two percent per year in the first 15 years of the transition.

The behavior of the interest rates and the investment-to-output ratios along the delayed transition is also consistent with the growth dynamics in the data, which are not easily accounted for by the standard growth theory (King and Rebelo, 1993). Unlike the neoclassical growth model where the interest rate and the investment rate start out at a very high level and decrease over time, our model generates rich dynamics for these variables, as we vary the degree of financial frictions and initial misallocation. For example, the real interest rate may remain nearly constant along the entire transition, and the investment rate may rise over time in the early stages of economic growth.

A central question raised by this research pertains to the sources and extent of misalloca-

[2]Financial frictions do have a large impact on the interest rate and the wealth distribution of the stationary equilibrium. As can be construed from the discussions above, more frictions translate into stronger ability-wealth correlation in the stationary equilibrium, reflecting entrepreneurs' self-financing.

tion in an economy. In our quantitative work, we consider initial misallocation of wealth and entrepreneurial talent that is consistent with the empirical evidence on generic distortions that impede the efficient allocation of factors across productive units (Hsieh and Klenow, 2007; Restuccia and Rogerson, 2007). In particular, we use the firm-level distortions in the Chinese and Indian manufacturing sector as measured by Hsieh and Klenow together with our model to infer the corresponding joint distribution of ability and wealth. Our quantitative analysis in this context suggests that financial frictions alone can only account for one sixth of the total distortions constituting the low TFP in developing countries—a re-statement of our first result that financial frictions only have small effects on the stationary equilibrium. In addition, our result on transition dynamics predicts that the removal of generic distortions will not instantaneously restore efficiency in the economy: The initial misallocation and financial frictions interact and delay the process of efficient reallocation of economic resources.

In summary, our results challenge and complement two commonly-held views in the literature. Our first result shows that the long-run effects of financial frictions are small, once realistic forward-looking saving motives are incorporated into quantitatively-oriented models. Our second result complements the recent literature that emphasizes the role of distortions and misallocation in lowering the steady-state TFP. We model one type of distortions (financial frictions) explicitly, and quantify their small effects on the steady-state TFP as well as their disproportionately large and persistent impact on the transition dynamics.

Related Literature We build on the theoretical literature that places financial frictions as the central issue on economic development.[3] We develop their ideas in ways that are empirically useful, by studying the transitional dynamics and the stationary equilibria of a broader class of quantitatively-oriented models with financial frictions.

Giné and Townsend (2004) and Jeong and Townsend (2005, 2007) have pioneered quantitative analysis for this class of models. They estimate and calibrate some models in this literature to the growth experience of Thailand. We share their interest in studying the role of financial frictions on transitional dynamics. However, we abstract from financial deepening which is the main driving force of their transition dynamics. Instead, we emphasize how the joint distribution of ability and wealth evolve endogenously over time under financial frictions, starting from an initial condition characterized by misallocation of economic resources.[4]

In related work, Amaral and Quintin (2005) and Caselli and Gennaioli (2005) find that credit market imperfections have quantitatively significant effects on the steady-state output and productivity. We recast their analysis in a setting that is more general in some important dimensions.

[3] Early contributions to this literature include Aghion and Bolton (1997), Banerjee and Newman (1993), Erosa (2001), Galor and Zeira (1993), Greenwood and Jovanovic (1990), Lloyd-Ellis and Bernhardt (2000), and Piketty (1997).

[4] More specifically, we incorporate into the model forward-looking endogenous saving decisions and heterogeneity in the returns to capital across entrepreneurs, both of which they abstract from. They do assume heterogeneity across individuals in the fixed setup cost of starting a business.

Caselli and Gennaioli use an exogenous constant saving-rate rule, which we replace with endogenous saving decisions. Unlike Amaral and Quintin who construct a two-period OLG economy, we solve a model with more realistic time horizons that does not lock in individuals to constraints for a long period of time: In a two-period OLG economy, if an individual is credit-constrained in one period, she is by construction credit-constrained for 15–20 years.

To quantify the effect of financial frictions and resource misallocation, one first needs to know the extent to which resources are misallocated in real-world economies. We follow the recent work of Hsieh and Klenow (2007) who, using Chinese and Indian manufacturing census data, show that misallocation of resources across productive units (plants) accounts for the low TFP of these economies. They also point out that, especially for China, the gradual reallocation of resources across firms over time is behind the recent aggregate productivity growth. Our paper is related to theirs in two ways: First, we use their evidence and methodology to calibrate the initial misallocation in our numerical exercises; Second, we provide a model that endogenizes the gradual reallocation of resources that they document.

Christiano (1989) and King and Rebelo (1993) point out that the neoclassical transitional dynamics is inconsistent with the observed growth experiences. They also study whether modified versions of the neoclassical growth model can account for the observed dynamics. The modifications include non-homothetic preferences, adjustment costs and a broader notion of capital, but all of them lead to some counterfactual implications for investment rates, interest rates and/or relative prices of installed capital and new investment goods. More recently, Chen et al. (2006) reconcile the neoclassical growth model with the post-war growth experience of Japan. They feed into the neoclassical model the measured realizations of the TFP path as an exogenous process, and show that the resulting dynamics is consistent with the data. In this context, we view our paper as an attempt at providing a theory of the TFP dynamics along the transitional paths based on the interaction of financial frictions and the initial misallocation of economic resources.[5]

More recently, the disappointing growth experiences of post-communist countries have motivated many researchers to study economic transitions. This literature focuses on the reallocation of factors from state to private enterprises, with a particular emphasis on worker flows and labor market frictions (Blanchard, 1997). Our contention is that capital and entrepreneurial talents were not appropriately aligned during the communist era, and that financial frictions delayed efficient reallocation of capital even after the liberalization.[6] Atkeson and Kehoe (1997) also attribute the delayed transition of these economies to misallocation of capital. In their model, capital cannot be swiftly reallocated across firms because it takes time for new private firms to accumulate complementary "organizational" capital.

A disparate literature in macroeconomics studies the stationary equilibria and transition dy-

[5]Similarly, adjustment costs could be thought of as a reduced form representation of financial frictions and resource misallocation, which are explicitly modeled in this paper.

[6]In the communist economies, the allocation of capital was as likely to be determined by the distribution of power as by productivity. See Blanchard (1997) and Roland (2000) and the references therein. Calvo and Coricelli (1992) argue that credit market frictions inhibited efficient reallocation of capital in Poland after the liberalization.

namics of related models featuring heterogeneity and financial frictions. Aiyagari (1994) shows how introducing uninsurable idiosyncratic risks leads to a larger aggregate capital stock and to a well-defined invariant distribution of wealth. Huggett (1997) studies the transition dynamics of Aiyagari's economy, but finds only small quantitative differences from those of a representative-agent model.[7] The case with aggregate shocks is studied by Krusell and Smith (1998). They show that, for the cases with idiosyncratic labor risk, a strong approximate aggregation result holds—that is, the distribution of wealth does not matter for aggregate dynamics in the stochastic stationary equilibrium.

With the issue of heterogeneity and transition dynamics seemingly resolved, some researchers have focused on the difficulty of these incomplete-market models in matching the highly-skewed wealth distribution in the US. More recently, Cagetti and De Nardi (2006) and Quadrini (2000) incorporate financial frictions into models with individual-specific technologies (entrepreneurship), and show that these elements explain the empirical wealth distribution. Intuitively, if there are financial frictions, highly-talented entrepreneurs will hold a large ownership stake in their own businesses, which translates into a fat right tail of the wealth distribution.[8] This literature primarily focuses on the wealth distribution of the stationary equilibria, and hence does not study the impact of financial frictions on the process of economic development.

Finally, the way we model financial frictions is also related to the macroeconomic literature on credit multipliers (Bernanke and Gertler, 1989; Bernanke et al., 1999; Kiyotaki and Moore, 1997). This literature focuses on how financial frictions transmit and propagate shocks at the business-cycle frequency, while our analysis pertains to longer-run economic phenomena.

1 Model

We study economies with individual-specific technologies and imperfect credit markets. Agents choose either to operate an individual-specific technology—i.e. become entrepreneurs, or to work for a wage. Individuals are heterogeneous with respect to their entrepreneurial ability and wealth. There is exogenous borrowing constraint that is proportional to an individual's net worth.

Heterogeneity and Demographics Individuals are heterogeneous with respect to their birth date t, their initial wealth a_t and their entrepreneurial ability $e \in \mathcal{E} = \{e_1, \cdots, e_N\}$. An individual's ability does not change over the course of her lifetime. Individuals face a constant probability of death $1 - \gamma$ each period, and it is assumed that this mortality risk is i.i.d. across individuals. When an individual dies, she is replaced by an off-spring who inherits her wealth, and partially

[7] Huggett does find qualitative differences between the models. For example, the transitional dynamics of aggregate capital in the heterogeneous-agent model is not necessarily monotonic.

[8] Entrepreneurship has implications for the level of aggregate capital stock as well. In models with idiosyncratic risks on the return to the individual-specific technologies, there are two opposing forces. The precautionary saving motive will push up the aggregate capital stock in the stationary equilibrium, while the uncertainty will discourage investment in the risky technology and hence capital accumulation. Angeletos (2007) works out the conditions for either force to prevail.

inherits her entrepreneurial ability. The intergenerational transmission of ability follows a Markov process with transition probability $\mu(e'|e)$, where e' is the ability of the off-spring and e is that of the parent. We denote by $\mu(e)$ the measure of type-e individuals in the invariant distribution. We denote by $G_t(e,a)$ the cumulative density function for the joint distribution of ability and wealth at the beginning of period t.

The population size is normalized to one, and there is no population growth.

Preferences Individuals discount their own future utility and the utility of their off-springs using the same discount factor β. The preferences over contingent plans for the consumption sequence of a dynasty from the point of view of an individual in period t are represented by the following expected utility function:

$$\mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} u(c_s).$$

Technologies In any given period, individuals can choose either to work for a wage or to operate an individual-specific technology. We label the latter option as entrepreneurship. We assume that an entrepreneur with talent e who uses k units of capital and hires l units of labor produces according to the following production function:

$$f(e,k,l),$$

which is assumed to be homogeneous of degree one, strictly increasing in all its arguments, and strictly concave in capital and labor, with $f(0,k,l) = 0$ and $\lim_{e\to\infty} f(e,k,l) = \infty$. In the appendix, as a robustness check, we also consider a non-convex production technology.

Credit Markets Productive capital is the only financial asset in the economy. Individuals can lend and borrow capital at interest rate r_t, subject to a quantity constraint. In particular, we exogenously limit borrowing in each period to a constant fraction $\lambda - 1 \geq 0$ of an individual's wealth at the beginning of the period. This constraint limits the capital usage of an entrepreneur to:

$$k_t \leq \lambda a_t.$$

In this context, λ measures the degree of credit frictions, with $\lambda = \infty$ corresponding to perfect credit markets. We choose this specification of credit frictions as it offers a parsimonious representation of a general prediction from various models of imperfect credit—the amount of borrowing is limited by the net worth of entrepreneurs (Buera, 2006; Evans and Jovanovic, 1989).[9]

[9] We abstract from modeling these collateral constraints endogenously, and instead focus on understanding the quantitative effects of credit frictions on aggregate dynamics. See Albuquerque and Hopenhayn (2004) and Clementi and Hopenhayn (2006) for recent examples deriving credit frictions from underlying contractual and informational frictions, and for discussions on their different implications.

Agents' problem The problem of an agent in period t can be written as:

$$\max_{\{c_s, a_{s+1}\}_{s=t}^{\infty}} \mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} u(c_s)$$
$$s.t.\ c_s + a_{s+1} \leq \max\{w_s, \pi(a_s; e_s, w_s, r_s)\} + (1+r_s)a_s,\ \forall s \geq t \tag{1}$$

where e_t, a_t and the sequence of wages and interest rates $\{w_s, r_s\}_{s=t}^{\infty}$ are given, and $\pi(a; e, w, r)$ is the profit from operating an individual technology. This indirect profit function is defined as:

$$\pi(a; e, w, r) = \max_{l,k \leq \lambda a} \{f(e, k, l) - wl - (\delta + r)k\}.$$

The input demand functions are denoted by $l(a; e, w, r)$ and $k(a; e, w, r)$.

A type-e individual with current wealth a will choose to be an entrepreneur if profits as an entrepreneur, $\pi(a; e, w, r)$, exceed income as a wage earner, w. This occupational choice can be represented by a simple policy function. Type-e individuals decide to be entrepreneurs if their current wealth a is higher than the threshold wealth $\underline{a}(e)$, where $\underline{a}(e)$ solves:

$$\pi(\underline{a}(e); e, w, r) = w.$$

For some e, there may not exist such an $\underline{a}$. In particular, if e is too low, then $\pi(a; e, w, r) < w$ for all a. In this case, this type of individuals will never become entrepreneurs. Intuitively, individuals of a given ability choose to become entrepreneurs if they are wealthy enough to run their businesses at a profitable scale. Similarly, agents of a given wealth choose to become entrepreneurs only if their ability is high enough.

If the wage and the interest rate are constant, $w_t = w$ and $r_t = r$ for all t, as is the case in a stationary equilibrium of this economy, then the recursive formulation of the sequence problem (1) is given by the following Bellman equation:

$$v(a; e) = \max_{c, a'} \{u(c) + \beta \mathbb{E}[v(a'; e') | e]\}$$
$$s.t.\ c + a' \leq \max\{w, \pi(a; e, w, r)\} + (1+r)a. \tag{2}$$

Competitive Equilibrium Given $G_0(e, a)$, a competitive equilibrium in this economy consists of sequences of joint distribution of ability and wealth $\{G_t(e, a)\}_{t=1}^{\infty}$, allocations $\{c_s(e_t, a_t), a_{s+1}(e_t, a_t), l_s(e_t, a_t), k_s(e_t, a_t)\}_{s=t}^{\infty}$ for all $t \geq 0$, and prices $\{w_t, r_t\}_{t=0}^{\infty}$ such that:

1. Given $\{w_t, r_t\}_{t=0}^{\infty}$, e_t, and a_t, $\{c_s(e_t, a_t), a_{s+1}(e_t, a_t), l_s(e_t, a_t), k_s(e_t, a_t)\}_{s=t}^{\infty}$ solves the agent's problem in (1) for all $t \geq 0$;
2. The labor and capital markets clear at all $t \geq 0$, which by Walras' law implies goods market clearing as well:

$$\sum_{e \in \mathcal{E}} \left[\int_{\underline{a}(e, w_t, r_t)}^{\infty} l(a; e, w_t, r_t) G_t(e, da) - G_t(e, \underline{a}(e, w_t, r_t)) \right] = 0,$$
$$\sum_{e \in \mathcal{E}} \left[\int_{\underline{a}(e, w_t, r_t)}^{\infty} k(a; e, w_t, r_t) G_t(e, da) - \int_0^{\infty} a G_t(e, da) \right] = 0,$$

3. The joint distribution of ability and wealth $\{G_t(e,a)\}_{t=1}^{\infty}$ evolves according to the equilibrium mapping:

$$G_{t+1}(e,a) = \gamma \int_{u \leq a} \int_{a'(e,v)=u} G_t(e,dv)\,du + (1-\gamma) \sum_{e_-} \mu(e|e_-) \int_{u \leq a} \int_{a'(e_-,v)=u} G_t(e,dv)\,du.$$

2 Quantitative Exploration

In this section, we explore the quantitative implications of the model for the long-run economy and for the transitional dynamics. We calibrate the stationary equilibrium of the benchmark economy to the US data on standard macroeconomic aggregates, firm-size distribution, firms' external financing, and income mobility/inequality. We first study how an economy's stationary equilibrium responds to different degrees of financial frictions. Then we analyze how financial frictions interact with initial resource misallocation and influence an economy's transition to its stationary equilibrium.

2.1 Calibration

We first describe the parametrization of the model, and then discuss the calibration strategy and results. For the sake of clarity, we choose a parsimonious parametrization that follows as much as possible the standard practices in the literature.

We choose a period utility function of the isoelastic form:

$$u(c) = \frac{c^{1-\sigma} - 1}{1-\sigma}.$$

We assume that an entrepreneur with talent e who hires k units of capital and l units of labor produces according to the following Cobb-Douglas production function:

$$f(e,k,l) = e^{\nu} \left(k^{\alpha} l^{1-\alpha}\right)^{1-\nu}, \tag{3}$$

where ν is the share of output going to the entrepreneur—$1-\nu$ is known as the span-of-control parameter (Lucas, 1978). Accordingly, $1-\nu$ represents the share of output going to the variable factors. Out of this, fraction α goes to capital, and $1-\alpha$ goes to labor.

The entrepreneurial ability e^{ν} is assumed to be a discretized version of an exponential distribution whose probability density is $\zeta \exp\{-\zeta e^{\nu}\}$ for $e^{\nu} \geq 1$. For intergenerational transmission of ability, we assume that an individual inherits the ability of her parent with probability ψ. With probability $1-\psi$, she draws, independently of her parent's ability, a new ability realization from the exponential distribution of ability given above. Obviously, ψ controls the persistence of ability across generations, while ζ determines the dispersion of ability in the population.[10]

[10] We discretize the support of the ability distribution into 30 grid points, with the minimum being 1.0 and the maximum, $e^{\nu}_{\max}$, being the 99.985-percentile of the exponential distribution parameterized by ζ. The second highest

	US data	Model	Parameter	Elasticity
Fraction of external financing	0.54	0.54	$\lambda = 5.0$	0.81
Top 10% employment	0.60	0.59		-1.49, -0.87
Entrepreneurs	0.08	0.08	$\nu = 0.17$, $\zeta = 8.8$	3.25, 1.50
Top 5% earnings	0.30	0.29		-0.72, -0.65
Earnings correlation across generations	0.40	0.40	$\psi = 0.35$	1.69
Interest rate (2 yr)	0.10	0.10	$\beta = 0.87$	-7.31

Table 1: Calibration

We now need to specify nine parameter values: two technological parameters α, ν, and the depreciation rate δ; two parameters describing the process for ability ψ and ζ; the degree of financial frictions λ; the subjective discount factor β and the reciprocal of the intertemporal elasticity of substitution σ; and the survival probability γ.

We let $\sigma = 1.5$ following the standard practice. A period in the model is set to two years, and we let $\gamma = 0.933$ so that the average duration of working lives is 30 years. The two-year depreciation rate is set at $\delta = 0.116$. We impose $\alpha(1-\nu) = 0.30$ to match the aggregate share of capital.

We are thus left with five parameters (λ, ν, ζ, ψ, and β). We calibrate them to match six relevant moments in the US data: the fraction of capital that is externally financed; the fraction of entrepreneurs; the employment share of the top decile of employers (establishments); the share of earnings generated by the top five-percentile; the persistence of earnings across generations; and the real interest rate over two years.

The second column of Table 1 shows the value of these moments in the US data. The fraction of externally-financed capital 0.54 is obtained by dividing the sum of credit market liabilities of the private sector from the Flow of Funds data by the aggregate capital stock of this sector from the NIPA accounts. In our model, entrepreneurs are owners of private businesses. We follow Cagetti and De Nardi (2006) and Castañeda et al. (2003) in identifying private entrepreneurs in the data as the self-employed and business owners. The largest—measured by employment—decile of establishments in the US account for 60 percent of the total employment. We target the earnings share of the top five-percentile (0.30) as reported by Cagetti and De Nardi (2006) and Castañeda et al. (2003). Solon (1999) surveys the estimates of the correlation between a child's and parents' "permanent" income in the US, and finds them to be between 0.3 and 0.5. We choose to match the mid-point of 0.4. Finally, as the target interest rate, we pick five percent per year, implying a 10 percent rate over a two-year period.

ability, e^{ν}_{29}, is the 98.8-percentile. The other 27 grid points are equi-distant between the minimum and the second highest ability. That is, the grid point corresponding to the j-th highest ability, $j = 1, \cdots, 29$, is given by $1+(j-1)d$, where $d = (e^{\nu}_{29}-1)/28$. The probability mass for each discretized ability type is $\exp\{-\zeta(e^{\nu}_{j}-1)\}-\exp\{-\zeta(e^{\nu}_{j}+d-1)\}$, for $j = 1, \cdots, 29$. The remaining probability is assigned to the maximum ability.

The third column of Table 1 shows the moments simulated from the calibrated model. Note that we are able to nail down this "over-identified" system of target moments with a judicious choice of parameter values. Even though in the model economy all six moments are jointly determined by the five parameters, each moment is primarily affected by one particular parameter. The last column of the table shows the elasticity of each moment to the corresponding parameter.

We briefly discuss the identification and the interpretation of some of the parameter values. The degree of financial frictions are inferred to be low—a relatively high value of $\lambda = 5.0$, mainly reflecting the large fraction of intermediated capital in the US economy. The span-of-control parameter $1 - \nu$ and the dispersion of ability ζ more or less jointly determine the earnings share of the top five-percentile and the fraction of entrepreneurs in the population, as well as the employment share of the largest decile of establishments. They are calibrated at $\nu = 0.17$ and $\zeta = 8.8$.[11] While both higher returns to scale (a lower ν) and more dispersion in the ability distribution (a lower ζ) imply more concentration in the distribution of employment and income, the returns to scale parameter has a relatively larger impact on employment. The parameter $\psi = 0.35$ leads to a model correlation of time-averaged earnings across generations close to 0.40. Finally, the model requires a two-year discount factor $\beta = 0.87$ to match the two-year interest rate of 10 percent.

2.2 Results for the Stationary Equilibrium

We first study the long-run effect of financial frictions. In Figure 1, we consider how allocations and prices of the stationary equilibria respond to changes in the collateral constraint parameter λ. The top left panel shows the effect of the collateral constraints on aggregate output and capital. Both variables are measured relative to their values in the benchmark ($\lambda = 5.0$). Financial frictions have only small effects on these variables. Even if we completely shut down financial intermediation, for example, the steady-state output declines by only 10 percent. This result is explained by the fact that a talented but poor individual will accumulate enough wealth quickly so that she can overcome the financial constraints and operate her productive technology at the maximal-profit scale: With high marginal returns to capital, the saving rate is correspondingly high. Therefore, in the stationary equilibrium, only a small fraction of individuals are bound by the constraints. Another factor mitigating the effect of frictions is the assumption that ability is positively correlated across generations. However, this only plays a minor role: Our result is robust to the assumption of i.i.d. ability across generations (not reported here). Self-financing, while potentially costly from an individual's perspective (Buera, 2006), is a good substitute to formal financial markets in the context of the aggregate economy.

We identify two modeling choices that make self-financing a viable substitute for formal credit markets here, unlike in the earlier literature. First, we assume that individuals have a realistic time horizon: One period in the model is equal to two years, and hence our individuals live through

[11] In the literature, the ν parameter is typically not calibrated but imposed. For example, Atkeson et al. (1996), Atkeson and Kehoe (2005), and Amaral and Quintin (2005) use $\nu = 0.15$, while Guner et al. (2006) choose 0.20. Cagetti and De Nardi (2006) do calibrate it at 0.10, based on the income and wealth distribution in the US.

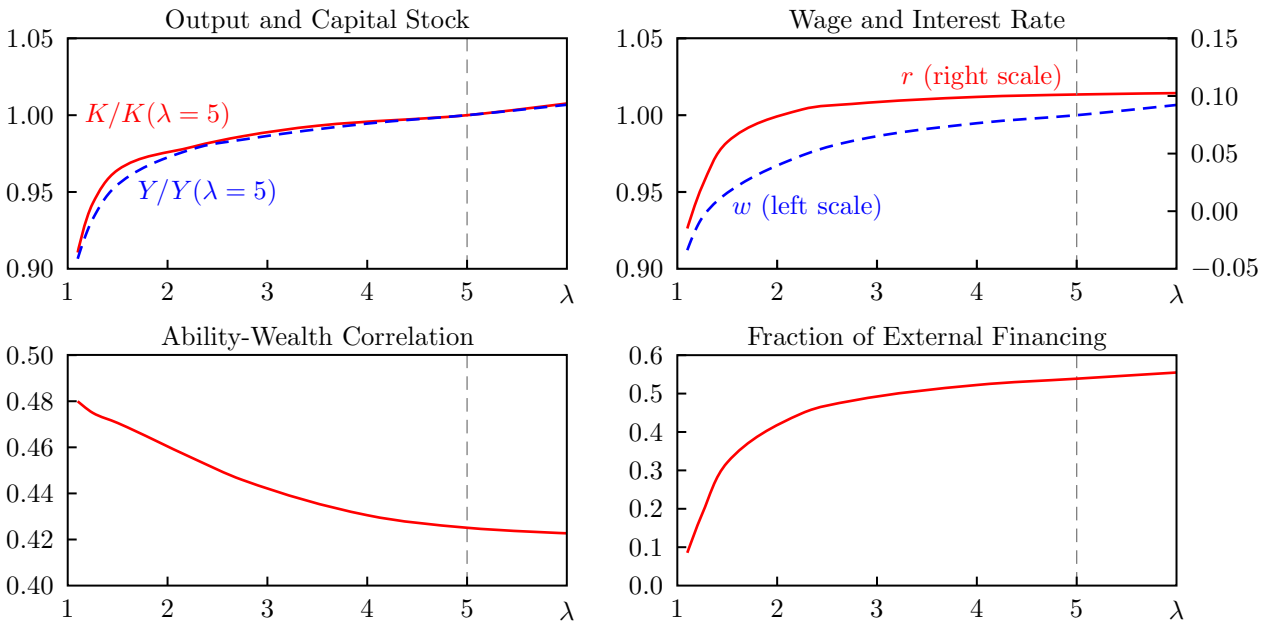

Fig. 1: Effect of Financial Frictions (λ) on the Stationary Equilibrium. We compute stationary equilibria that correspond to various values of λ, while holding all other parameters fixed as in Table 1. Six moments of interest are plotted against the degree of financial frictions, λ. Output, capital stock and wage are normalized by their counterparts in the benchmark stationary equilibrium with $\lambda = 5$, which is represented by the dashed vertical lines. Note that a lower λ implies more financial frictions.

many periods with the option of adjusting their behavior every period. By comparison, in a two-period OLG economy, if a talented individual is born poor, she is by construction condemned to credit constraints for at least half her lifetime. Second, individuals in our model are forward-looking and make saving decisions to maximize their objective function. One implication is that talented individuals, especially when they are poor (and credit-constrained), exhibit much higher saving rates than those who are not as talented. Note that earlier theoretical literature typically assumes exogenous myopic saving decision rules.

One period is equal to	Survival prob. γ	Discount factor β	Depreciation δ	Relative steady-state output $Y(\lambda = 1.25)/Y(\lambda = 5)$
2 years	0.93	0.87	0.12	0.93
5 years	0.83	0.71	0.27	0.87
15 years	0.50	0.35	0.61	0.78

Table 2: Individuals' Horizon and the Effect of Financial Frictions on Steady-State Output

To better understand the small long-run effect of financial frictions in this context, we re-do our numerical exercise with different lengths of one period in the model. Table 2 reports how the marginal effect on the steady-state per-capita output of financial frictions—decreasing λ from 5.0 to 1.25—changes, when we consider model economies where one period is equal to 5 or 15 years. As we change the length of one period in the model, we appropriately adjust the survival probability,

discount factor and depreciation rate. As we reduce the number of periods in which individuals can adjust their wealth—or, put differently, lengthen the time horizon over which individuals are locked in to constraints, the steady-state impact of credit frictions gets larger. With one period in the model equal to 15 years, a high-ability individual born poor is condemned by construction to binding credit constraints for at least 15 years. She is unlikely to live through enough periods (multiple of 15 years) to save up and overcome the financial frictions. Thus, in the steady state, a larger fraction of individuals will be credit-constrained than in the case where one period is defined as two or five years.[12]

We also explore the robustness of our result to the specification of the individual technology. In the appendix, we consider a technology with minimum scale, which is potentially useful in accounting for the observed magnitude of heterogeneity in returns to scale across firms in poor countries (Banerjee and Duflo, 2005). Our result that financial frictions alone have only small effects on the steady state remains robust to this alternative technology specification.

Note that credit market imperfections have a more pronounced impact on the equilibrium interest rate, ability-wealth correlation (reflecting self-financing by talented entrepreneurs) and, of course, the share of aggregate capital that is externally financed (Figure 1).

2.3 Results for the Transitional Dynamics

We now study how financial frictions affect the transition dynamics of the model. In particular, we study the transition dynamics of economies that start with a low level of aggregate capital stock and with different degrees of misallocation as measured by the correlation of entrepreneurial ability and wealth.[13] The initial conditions in our quantitative exercise represent an economy's steady state under generic distortions that impede efficient allocation of resources. Such distortions are considered important in accounting for the observed resource misallocation in less-developed countries (Hsieh and Klenow, 2007; Restuccia and Rogerson, 2007). In this context, the transitional dynamics that we study here provide a theory of how an economy will evolve under financial frictions once all other distortions are removed.

With initial conditions calibrated to empirical evidence on misallocation, we show that financial frictions have a large impact on the transitional dynamics. In fact, financial frictions and capital misallocation are keys to understanding the observed economic transitions that are not easily explained by the standard neoclassical model.

To illustrate how misallocation and financial frictions jointly determine transition dynamics,

[12] This exercise explains why our result differs from that of Amaral and Quintin (2005) or Caselli and Gennaioli (2005), who find quantitatively significant effects of credit market imperfections on aggregate output and productivity in the steady state. Caselli and Gennaioli assume one-period-lived individuals and exogenous constant saving-rate rule, while Amaral and Quintin study a two-period OLG economy. Our numerical results (not reported here) suggest that the assumption of exogenous constant saving-rate decision rule also tends to make the effect of financial frictions bigger.

[13] Unlike in the neoclassical growth model where the aggregate capital stock (unconditional first moments) is the sufficient statistic for the aggregate dynamics, in our model the aggregate dynamics are a function of the entire joint distribution of entrepreneurial ability and wealth.

we take two economies with different initial conditions, and work out their evolution under two different degrees of financial frictions, $\lambda = 5.0$ and 1.5. While $\lambda = 5.0$ is our calibration for the US, $\lambda = 1.5$ is a number consistent with the external financing data of a typical less-developed country. The aggregate capital stock in these two initial conditions is half that of the $\lambda = 5.0$ stationary equilibrium. In the first economy (dashed line, Figure 2), the initial distribution is constructed by halving the support of the benchmark wealth distribution, and hence the correlation between initial wealth and entrepreneurial ability (e^ν) is 0.43. The ability-wealth correlation for the second economy (solid line, Figure 2) is 0.07. That is, the second economy has more misallocation of initial resources than the first one does. In constructing this second initial condition, we use the distortions in developing economies as measured by Hsieh and Klenow (2007). In this regard, the initial misallocation we consider here is empirically relevant. A more detailed discussion of how we apply Hsieh and Klenow's method is given in Section 2.4.

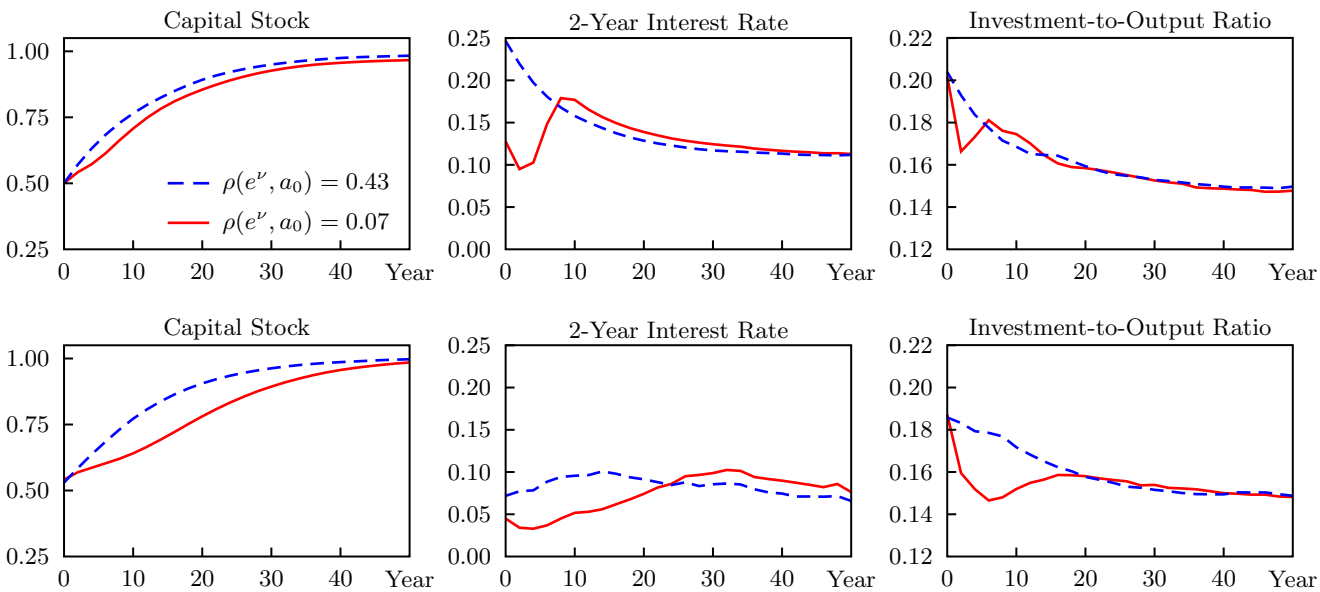

Fig. 2: Transition Dynamics with Financial Frictions. There are two initial ability-wealth distributions. The first (dashed line) is obtained by halving the support of the benchmark wealth distribution, preserving the ability-wealth correlation of $\rho(e^\nu, a_0) = 0.43$. The second (solid line) has more resource misallocation, with $\rho(e^\nu, a_0) = 0.07$. The top row depicts the evolution of the two economies—i.e. initial distributions—under $\lambda = 5.0$. The bottom panels plot the transition under more severe financial frictions, $\lambda = 1.5$. The aggregate capital stock is normalized by that in the respective stationary equilibrium.

Slow Convergence and Investment/Interest Rate Dynamics The top and bottom left panels in Figure 2 show the aggregate capital for these two economies over time.[14] The top panel is with $\lambda = 5.0$ (US benchmark) and the bottom is with $\lambda = 1.5$ (more frictions). In the second economy (solid line, top and bottom panels), with more misallocation of initial resources, only a

[14]The actual computation is done for $T = 125$ periods (250 years), but we only show the first 50 years of transition.

certain fraction of the misallocated capital can be reallocated instantaneously to more productive entrepreneurs because of the credit constraints. The misallocation can be undone only over time, as the higher-ability types accumulate wealth and the low-ability types retire from entrepreneurial activities. Therefore, for a prolonged period of time, the second economy is less productive than the first (dashed line), and it converges to the steady state at a slower pace. The effect of misallocation is more pronounced when the financial frictions are more severe ($\lambda = 1.5$, bottom panel): Compared to the first economy (dashed line), the second (solid line) takes more than twice as long to cover half the distance between the initial aggregate capital stock and the stationary equilibrium capital stock. In the early stages of the second economy's transition, a larger fraction of resources are controlled by low-ability types than in the first one. Given that low-ability types' return to capital is lower than high-ability types', aggregate output and saving are lower. This explains both the slower transition to the steady state and the lower investment rate in the second economy.

The main point of King and Rebelo (1993) is that the neoclassical model's prediction on interest rates is inconsistent with data. The plots show that, with financial frictions and misallocation of initial resources, we generate interest rate paths that start low and go up gradually over time. In the early stages, the demand for capital is restricted by the borrowing constraint—the poor high-ability types cannot use much capital because they have negligible net worth. This drives down the market interest rate. As high-ability types become richer, they can use more capital, and the market interest rate rises.

In summary, with financial frictions and misallocation of initial resources, our model generates slow growth, low interest and investment rates in the early stages of economic development. These features are broadly consistent with the transition dynamics of the post-communist economies and of the miracle economies. More importantly, as we can see from comparing the top with the bottom panels of Figure 2, financial frictions determine how persistent the effect of initial misallocation is—by controlling the speed with which resources get reallocated for more productive use.

Endogenous TFP Dynamics Another feature of our model economies is the endogenous TFP dynamics. Under $\lambda = 5.0$ and 1.5, we plot the output and the imputed TFP of the two economies with different initial conditions (Figure 3). Even though the two economies start with the same aggregate capital stock, the second economy ($\rho(e^{\nu}, a_0) = 0.07$, solid line) initially produces much less than the first one (dashed line), because the financial frictions limit instantaneous reallocation of resources for more productive use. The TFP in these economies captures the effect of misallocation and financial frictions. With $\lambda = 1.5$ (bottom panels), as the initial misallocation is unwound over time for the second economy (solid line), the imputed TFP rises at about two percent per year for the first 15 years, although there is no change on the technology side.

Again, by comparing the top with the bottom panels, we see that the same initial misallocation has larger and more persistent effects when there are more financial frictions ($\lambda = 1.5$).

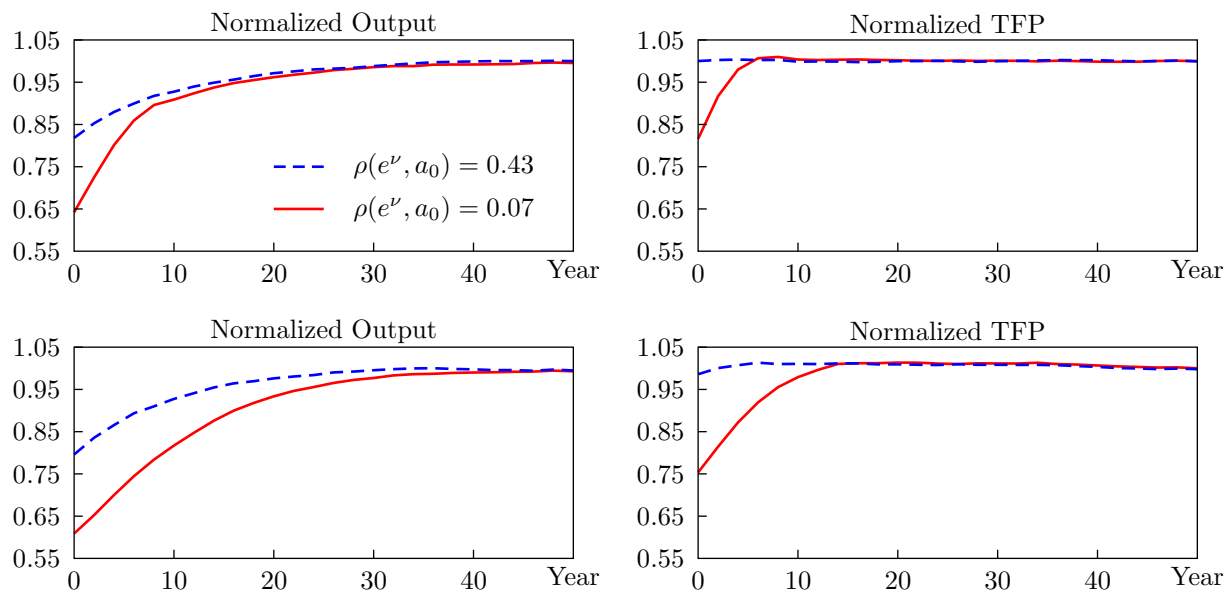

Fig. 3: Output and TFP with Misallocation. With $\lambda = 5.0$ (top panels) and $\lambda = 1.5$ (bottom panels), we plot the aggregate output and the imputed TFP—calculated using the standard growth accounting—for two initial distributions. The first (dashed line) is again obtained by halving the benchmark wealth distribution ($\rho(e^\nu, a_0) = 0.43$). The other (solid line) has more resource misallocation, with $\rho(e^\nu, a_0) = 0.07$. Both the output and the imputed TFP are normalized by their levels in the respective stationary equilibrium.

2.4 Constructing Initial Conditions: Evidence on Misallocation

We have shown that the transitional dynamics in economies with misallocation of initial resources are significantly different from what is predicted by the neoclassical growth model. Are these departures from the standard neoclassical theory empirically relevant? In other words, do actual economies exhibit misallocation of wealth and entrepreneurial ability? We have noted that our initial misallocation is consistent with the empirical evidences of Hsieh and Klenow (2007) on the misallocation of labor and capital in developing countries. Here we provide a more detailed description. We first compare the misallocation in the steady state of our model to their measures of distortions, and quantify the fraction of the measured distortions that can be accounted for by financial frictions alone. We then use their data on distortions together with our model to infer the correlation of wealth and ability in an economy characterized by distortions.

We first provide a brief review of the findings in Hsieh and Klenow (2007). Using plant-level information on capital and labor inputs and value added, they calculate output and capital wedges $\tau_{y,i}$ and $\tau_{k,i}$ that rationalize the observed allocation of factors:

$$\{k_i, l_i\} \equiv \arg\max\,(1-\tau_{y,i})\, e_i^\nu (k_i^\alpha l_i^{1-\alpha})^{1-\nu} - w l_i - (1+\tau_{k,i})(r+\delta)k_i,$$

where i indexes plants. They summarize the information embedded in the wedges by reporting the distribution of the log deviation from the mean of $\frac{(1+\tau_{k,i})^\alpha}{(1-\tau_{y,i})}$. We hereafter refer to this ratio as the

measure of misallocation.[15]

They find that, compared to the US economy, the misallocation of factors across plants in China and India are much more prevalent: The standard deviation of the misallocation measure is 0.74 for China in 1998 and 0.79 for India in 1994, while it is 0.42 for the US in 1997. The wedges in China and India are also more positively correlated with the productivity of plants—implying that more productive plants face more distortions—than those in the US: 0.49 for China in 1998 and 0.37 for India in 1994, compared to 0.18 in the US in 1997.

To put into perspective our finding that financial frictions have only small effects in the steady state, we calculate the capital wedges τ_k in the stationary equilibrium with $\lambda = 1.5$. The standard deviation of the measure of misallocation is 0.14 and its correlation with plant productivity is 0.30. Thus, financial frictions alone can explain only one sixth of observed misallocation.

Now we describe how we use the measured misallocation of Hsieh and Klenow to construct empirically-relevant initial ability-wealth misallocation. We consider an economy where entrepreneurs face large idiosyncratic output distortions $\tau_{y,i}$, implying that an individual's perceived productivity as an entrepreneur is $(1-\tau_{y,i})\, e_i^{\nu}$, where e_i^{ν} is her true ability. Using the estimated distribution of the output distortions provided by Hsieh and Klenow, we generate a joint stationary distribution of wealth and this perceived ability under financial frictions. From this stationary distribution, we can reverse-engineer the joint distribution of wealth and the true ability e_i^{ν}. With $\lambda = 1.5$, the correlation between wealth and the true entrepreneurial ability thus constructed is 0.07, while that between wealth and the perceived ability is 0.47.

2.5 Aggregate Impact of Ability-Wealth Distribution

To better understand how misallocation affects the transitional dynamics, we study what particular dimensions of wealth distribution matter for aggregate output.

With perfect credit markets, this economy is isomorphic to a standard neoclassical growth model with the aggregate production function given by:

$$F(K) = \max_{e_m, 0 \le \iota \le \mu(e_m)} f\Big(\sum_{e > e_m} \mu(e) e + \iota e_m, K, \sum_{e < e_m} \mu(e) + \mu(e_m) - \iota\Big)$$

where K is the aggregate capital stock. Furthermore, if the individual technology is of the Cobb-Douglas form, $f(e,k,l) = e^{\nu}\left(k^{\alpha} l^{1-\alpha}\right)^{1-\nu}$, the aggregate production function simplifies to:

$$F(K) = A(\mu)\, K^{\alpha(1-\nu)}, \tag{4}$$

$$A(\mu) = \max_{e_m, 0 \le \iota \le \mu(e_m)} \left(\sum_{e > e_m} \mu(e)\, e + \iota e_m\right)^{\nu} \left(\sum_{e < e_m} \mu(e) + \mu(e_m) - \iota\right)^{(1-\nu)(1-\alpha)},$$

where $A(\mu)$ embodies the effect of the distribution of entrepreneurial ability on aggregate output. Behind the aggregate production function lie optimal allocations of individuals to occupations—workers and entrepreneurs—and of capital and labor to active entrepreneurs.

[15] The ratio can be re-written as $\frac{(1+\tau_{k,i})^{\alpha}}{1-\tau_{y,i}} = \left(\frac{1}{MPK_i}\right)^{\alpha} \left(\frac{1}{MPL_i}\right)^{1-\alpha}$. That is, it provides a measure of the dispersion of the marginal products of capital and labor across plants, weighted by the output elasticities.

With financial frictions, however, aggregate output is not a simple function of aggregate factors of production as in Equation (4). Instead, aggregate output is now a function of the joint distribution of wealth and entrepreneurial ability:

$$Y = \tilde{F}\left(G\left(e, a\right)\right). \tag{5}$$

While (4) is a proper production function—one that prescribes the maximum output that can be obtained with a given amount of capital and labor, Equation (5) is an equilibrium object describing the aggregate output in an economy with a given distribution of resources and a market structure.

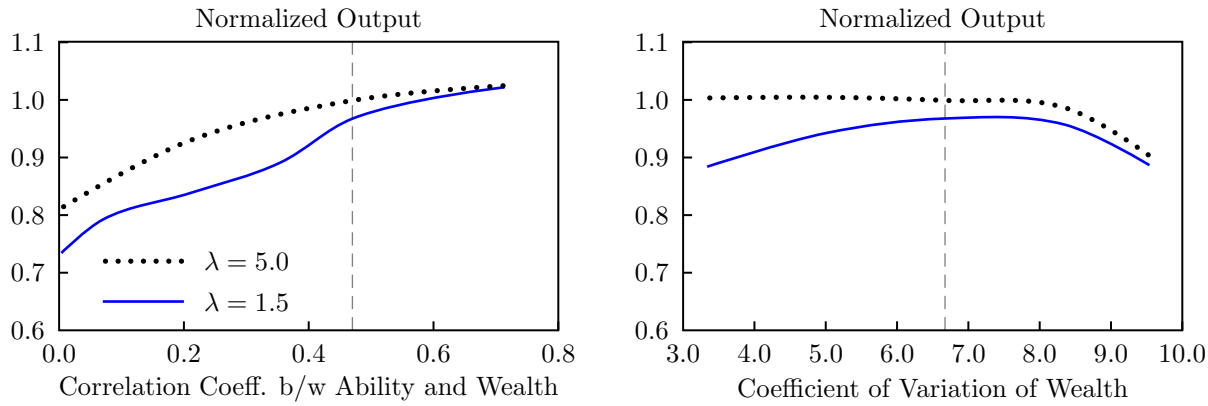

Fig. 4: Aggregate Output as a Function of the Joint Distribution of Wealth and Talent. For various ability-wealth distributions, we compute the aggregate output in the corresponding *static* equilibria under $\lambda = 1.5$ (solid lines) and $\lambda = 5.0$ (dotted lines). The output is then normalized by that of the benchmark distribution under $\lambda = 5.0$. The left panel plots aggregate output against the coefficient of correlation between ability (e^{ν}) and wealth of the given distribution. These distributions have the same marginal distribution of ability and wealth. The correlation coefficient of ability and wealth in the benchmark equilibrium is 0.47, which is denoted with a dashed vertical line. In the right panel, we vary the wealth variance while holding its mean and correlation with ability fixed. The coefficient of variation of wealth in the benchmark distribution is 6.67, denoted by a dashed line.

Figure 4 represents Equation (5) by showing how equilibrium output behaves as a function of the correlation of wealth and entrepreneurial ability (left panel), and the variance of wealth (right panel). A point on the horizontal axis corresponds to a particular ability-wealth distribution. The dashed vertical lines correspond to the correlation and variance in our benchmark distribution for this exercise.[16] For a given joint distribution of ability and wealth, we solve for the *static* competitive equilibrium with two different degrees of financial friction: $\lambda = 5.0$ and $\lambda = 1.5$.

In the left panel, starting with the benchmark distribution that has an ability-wealth correlation coefficient of 0.47, we vary the ability-wealth correlation while keeping the marginal distributions of wealth and ability unchanged. It can be seen that aggregate output is sensitive to the correlation between wealth and ability.[17] More misallocation (lower ability-wealth correlation) leads to lower

[16] Here our benchmark distribution is the invariant distribution with $\lambda = 1.5$ constructed in Section 2.2.

[17] We obtain distributions with lower correlation of wealth and ability by randomly reallocating the wealth and ability for a subset of individuals. With more reallocation, the correlation of wealth and ability gets lower. To obtain distributions with a larger correlation, we mix the original distribution of wealth with a distribution where each agent can operate the technology at her respective efficient scale.

output, because the financial frictions hinder efficient reallocation of resources. With more frictions ($\lambda = 1.5$, solid line), output is even lower.

In the right panel, starting from the benchmark distribution, we vary the unconditional variance of wealth while keeping the ability-wealth correlation and the marginal distribution of ability unchanged. Aggregate output is not affected as much, implying that what matters most for aggregate output is the degree of ability-wealth misallocation, measured here by their correlation coefficient.

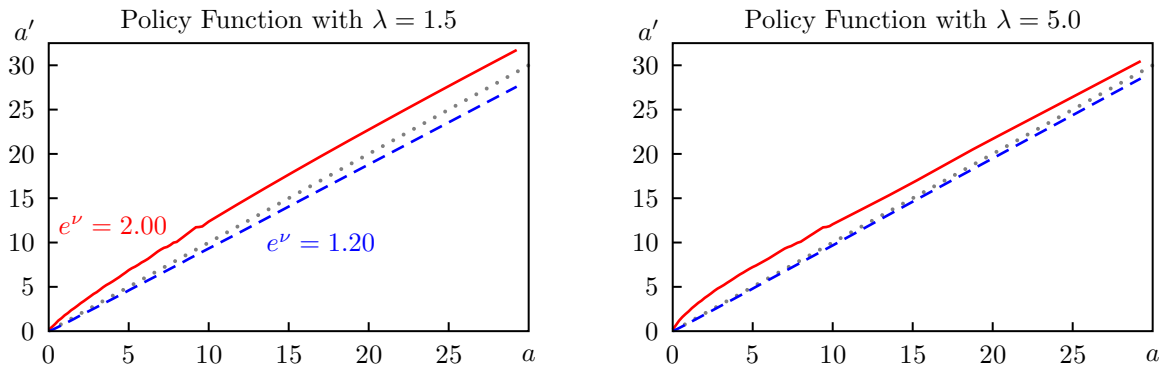

Fig. 5: Policy Functions in Stationary Equilibria. We plot individuals' choice of the next period asset holdings against their current holdings in stationary equilibria. Individuals with different entrepreneurial abilities have different policy functions. The solid lines correspond to individuals with the highest ability, while the dashed lines are for the median ability individuals. The 45-degree line is delineated with dots. Left panel is for the stationary equilibrium with $\lambda = 1.5$, and the right for $\lambda = 5.0$.

While Figure 4 underlines the *static* interaction between financial frictions and ability-wealth distribution, Figure 5 shows their *dynamic* interactions. The solid lines are the policy function (next period's asset holdings as a function of current wealth) of the most able entrepreneurs ($e^{\nu}_{\max}$). The dashed lines correspond to that of the median entrepreneurial ability types. These are policy functions in the stationary equilibrium under $\lambda = 1.5$ (left panel) and under $\lambda = 5.0$ (right panel). The first thing to note is that different ability types have markedly different policy functions. While the median ability type mostly dissave, the high ability type is keen on accumulating wealth—especially so when poor, as can be seen from the slope of the policy function. Clearly, it is important to know how the aggregate capital stock is distributed across different ability types, if one wants to predict the aggregate dynamics. With a lower degree of financial frictions (right panel), the difference in the policy functions across ability types diminishes. In the extreme case of perfect credit markets, all the types will have the same policy function, and hence how wealth is distributed will not matter for aggregate dynamics.

Figures 4 and 5 are a preview of a result that will be highlighted in the next section: For the aggregate economy, wealth inequality matters to the extent that it reflects the allocation of wealth across ability types.[18]

[18] For a simple heuristic argument, consider the case where the policy functions conditional on ability type e are approximately linear, as is the case in Krusell and Smith (1998): $a' = s_0(e) + s_1(e)\,a$. Aggregating over ability-wealth

3 Heterogeneity and Aggregate Dynamics

The previous section shows that the initial distributions of wealth and talent have quantitatively significant effects on the transitional dynamics. On the other hand, differences in the allocation of wealth *within* ability types are found to have smaller effects (Figure 4). These comparative statics do not exhaust the potentially complex interactions between the wealth distribution (an infinite-dimensional object) and aggregate dynamics.[19] In this section, we characterize which aspects of the wealth and talent distribution are important for explaining the aggregate dynamics of the economy. In particular, we provide an answer to the following question: Which is the minimum set of moments of the wealth and talent distribution that will suffice for a "good" approximation of the law of motion for aggregate variables?

To be more specific, let $G_t(e,a)$ be the joint distribution of wealth and entrepreneurial ability in period t, and $m_t \in R^M$ be a vector with M moments of $G_t(e,a)$.[20] We are interested in characterizing the smallest set of moments m_t that can closely approximate the evolution of prices and of themselves in a Markovian sense, starting from a wide range of initial wealth distributions $G_0(e,a)$.

Using the algorithm described in the appendix, we solve for the transitional dynamics of $I = 12$ economies that only differ in their initial distribution of wealth, $G_0(e,a)$. This gives us I histories of length T for the joint ability-wealth distribution, $G_t(e,a)$. We then find, within a given class of functions, the function H that best fits the evolution of aggregate variables—including prices—for all of the I histories, i.e., $w_t = H^w(m_t)$, $r_t = H^r(m_t)$, and $m_t = H(m_{t-1})$.[21] For example, if aggregate capital is the only moment considered, the approximation is given by: $w_t = H^w(K_t)$, $r_t = H^r(K_t)$, $K_t = H(K_{t-1})$.[22]

We find that, for our model, first and second moments of wealth conditional on occupational choices are sufficient statistics for "good" approximation of the law of motion for aggregate quantities and prices.

Table 3 illustrates this point. It shows the root mean squared errors from approximating the evolution of prices and aggregate quantities in period t as functions of a given set of moments. In Column (1) of Table 3 we show the root mean squared errors of approximating prices and the first

types, we obtain a law of motion for the aggregate wealth:

$$A' = \bar{s}_0 + \sum_e s_1(e)\,\mathbb{E}(a|e)\,\mu(e),$$

where $\bar{s}_0 = \sum_e s_0(e)\,\mu(e)$. Thus, as long as $s_1(e)$ is different across ability types, conditional means are important for describing the evolution of aggregate wealth.

[19] The current distribution of wealth might not fully encapsulate the entire history of the economy. See Duffie et al. (1994) and Miao (2006) for a discussion on this topic. Certainly, the algorithm we use to compute the transitional dynamics does not restrict us to recursive competitive equilibria.

[20] Throughout the exercises, we fix the marginal distribution of entrepreneurial talent in the population in any period to its invariant distribution.

[21] We use tensor products of Chebyshev polynomials for functional approximation.

[22] Notice that, in Krusell and Smith (1998), prices are a function of aggregate capital only, and the evolution of the log of aggregate capital can be well approximated as a linear AR(1) process.

Object of approximation	(1) RMSE	(2)	(3)	(4)	(5)
		RMSE relative to (1)			
Interest rate	0.055	36.0%	59.3%	4.9%	10.3%
Wage	0.135	31.6%	60.8%	8.2%	12.3%
Aggregate capital	0.056	30.6%	24.3%	4.3%	5.9%
Standard deviation of wealth	0.366	6.3%	44.1%	1.0%	1.7%

1. The columns correspond to approximations based on different sets of moments included in m_t. (1): the mean of wealth only; (2): the mean and the standard deviation of wealth; (3): the mean wealth and the average wealth of entrepreneurs; (4): all the moments included in (2) and the ability-wealth correlation coefficient; (5): all the moments in (2) plus the mean and the standard deviation of the wealth of entrepreneurs. We approximate the first 40 periods (80 years) of the transitional dynamics for 12 different initial distributions of wealth. We use tensor products of Chebyshev polynomials of orders 38, 7, 4 and 3 respectively for (2), (3), (4), and (5). The number of the Chebyshev regressors are 39, 49, 64 and 81.

2. The unit of the root mean squared errors (1) for the approximation of the interest rate is percentage points. The units of the root mean squared errors for the approximation of the other variables are log deviations.

Table 3: Approximation Errors

two unconditional moments of the wealth distribution solely as functions of aggregate capital. As suggested by the results describe in Section 2.3, aggregate capital is not a sufficient statistic for the aggregate dynamics. This is particularly true for interest rates, whose root mean squared error is 5.5 *percentage points*.

Information on the distribution of wealth across agents are important in describing the aggregate dynamics. Moreover, as shown in Column (2) of Table 3, higher-order unconditional moments are not informative enough. The approximation errors of predicting prices are reduced by less than two thirds. Intuitively, more unconditional inequality could be associated with either more or less misallocation of capital. Information about the wealth held by entrepreneurs can better predict the evolution of prices, aggregate capital, and inequality. Indeed, by incorporating a measure of the unconditional inequality and the mean and the dispersion of entrepreneurs' wealth—Column (5), we reduce the approximation error by 90 percent relative to the case where only aggregate capital is used. The unconditional variance captures the inequality both across types and within an ability type, while the mean entrepreneurial wealth measures inequality across types. The approximation error is very close to the one from using the correlation of wealth and ability, a more direct but harder-to-observe measure of the allocation of wealth and ability—Column (4). In summary, a good approximation only requires a small set of moments providing information about the extent of the ability-wealth allocation.

To take stock of things, unlike in Krusell and Smith (1998), in our model economy heterogeneity matters for aggregate dynamics.[23] Still, we recover a partial counterpart of the approximate

[23] One important distinction is that Krusell and Smith obtain the approximate aggregation result for stochastic

aggregation result. Conditional first and second moments are sufficient state variables for aggregate dynamics.

4 Concluding Remarks

In this paper, we incorporate financial frictions and entrepreneurship into an otherwise-standard neoclassical growth model, and use a calibrated version of the model to quantify the role of financial frictions in economic development. We find that financial frictions have small effects on the long-run output per capita. However, financial frictions do have a large impact along the transition to the steady state, especially when capital is misallocated initially. Our model economy converges slowly to the steady state, with the interest rate, investment rate and TFP starting low and rising over time.

We view this paper as a first step in building quantitative models to better understand the dynamics of development after growth-enhancing reforms. One commonly-held view on the income differences across countries is that there are barriers to adopting more productive modern technologies in poor countries (Hall and Jones, 1999; Parente and Prescott, 2000). Our model can be applied to study what will happen when such barriers are ratcheted down. The magnitude and the speed of capital reallocation from traditional sectors to modern manufacturing sectors will be determined by the existing resource misallocation across sectors—barriers inevitably imply misallocation—and the degree of financial frictions. In this context, our work is also complementary to the literature that explains cross-country income differences with institutions (Acemoglu et al., 2005). In particular, our result on transition dynamics predicts that the adverse impact of inefficient institutions will outlast them by decades. The post-communist transition of Eastern Europe is a relevant example, given the rampant resource misallocation during the communist era and the abrupt liberalizations that followed. For another example, many Latin American economies' disappointing performance after market-oriented reforms in the 1990s (Cole et al., 2005; Morley et al., 1999; Mukand and Rodrik, 2005) can be partly explained by the slow reallocation of capital toward more efficient technologies: Unequal distribution of resources has been a perennial hallmark of Latin America (Deininger and Squire, 1998; Sokoloff and Engerman, 2000).

stationary equilibria, while we only analyze the transition paths of an economy without aggregate shocks.

Appendix

A Numerical Algorithm

A.1 Computing the stationary equilibrium

Stationary Competitive Equilibrium A stationary competitive equilibrium in this economy consists of an invariant joint distribution of ability and wealth $G_\infty(e,a)$, policy functions $c(e,a)$, $a'(e,a)$, $l(e,a)$, $k(e,a)$, and prices w, r such that:

1. Given w and r, $c(e,a)$, $a'(e,a)$, $l(e,a)$, $k(e,a)$ solve the agents' problem (2);
2. The labor and credit markets clear, which by Walras' law implies goods market clearing as well:

$$\sum_{e\in\mathcal{E}}\left[\int_{\underline{a}(e,w,r)}^{\infty} l(a;e,w,r)\,G_\infty(e,da) - G_\infty(e,\underline{a}(e,w,r))\right] = 0,$$
$$\sum_{e\in\mathcal{E}}\left[\int_{\underline{a}(e,w,r)}^{\infty} k(a;e,w,r)\,G_\infty(e,da) - \int_0^\infty aG_\infty(e,da)\right] = 0;$$

3. The stationary joint distribution of ability and wealth $G_\infty(e,a)$ solves:

$$G_\infty(e,a) = \gamma\int_{u\le a}\int_{a'(e,v)=u} G_\infty(e,dv)\,du + (1-\gamma)\sum_{e_-}\mu(e|e_-)\int_{u\le a}\int_{a'(e_-,v)=u} G_\infty(e,dv)\,du.$$

We solve for the stationary equilibrium of this economy based on the nested fixed-point algorithm of Aiyagari (1994). The difference is that we have to iterate on both wage w and interest rate r until both labor and capital markets clear in the stationary equilibrium. We start by fixing a T, which is the period by which the economy must have reached the steady state. We choose T to be 250 (500 years). We numerically verify that increasing T to 300 has virtually no effect on the invariant distribution.

1. Guess the interest rate in the invariant distribution, r^i.
2. Guess the wage in the invariant distribution, $w^{i,j}$.
3. Given the guesses on interest rate and wage, solve the individuals' problem in the stationary equilibrium—Problem (2). Given the optimal decision rule, simulate N individuals for T periods. We set $N = 60,000$.
4. Check the labor market clearing condition in period T. If there is excess labor demand (supply), choose a new wage $w^{i,j+1}$ that is greater (smaller) than $w^{i,j}$.
5. Repeat Steps 3–4 until the labor market clears in period T.
6. Check the capital market clearing condition in period T. If there is excess capital demand (supply), choose a new interest rate r^{i+1} that is greater (smaller) than r^i.
7. Repeat Steps 2–6 until the capital market also clears in period T.

A.2 Computing the transition dynamics

To compute the entire transition dynamics, we have to iterate on the wage and interest rate sequences. Taking the wage and interest rate sequences as given, we solve for the individuals' problem—Problem (1), and then check whether labor and capital markets clear for all periods. We fix T at 125 (250 years). We numerically verify that increasing T to 150 has virtually no effect on the transition dynamics.

1. Guess at an interest rate sequence $\{r_t^i\}_{t=0}^T$.
2. Guess a wage sequence $\{w_t^{i,j}\}_{t=0}^T$. Compute the value function of the stationary equilibrium, and let $v_T(a;e) = v(a;e)$. By backward induction, taking the wage sequence $\{w_t^{i,j}\}_{t=0}^T$ and the interest rate sequence $\{r_t^i\}_{t=0}^T$ as given, compute the value function $v_t(a;e)$ for $t = T-1, \ldots, 0$. Using the optimal decision rule, simulate N individuals for T periods. We again set $N = 60,000$. Check whether the labor market clears in every period. Taking the individuals' capital holdings as given, construct a sequence $\{\varpi_t^{i,j}\}_{t=0}^T$ that clears the labor market for each period. Update the wage sequence: $w_t^{i,j+1} = \eta_w \varpi_t^{i,j} + (1-\eta_w) w_t^{i,j}$, $\forall t$, with $\eta_w \in (0,1)$. Iterate on the wage sequence until convergence.
3. Once the wage sequence converges, check whether the capital market clears in all periods. Taking the individuals' capital holdings as given, construct a sequence $\{\iota_t\}_{t=0}^T$ that clears the static capital rental market for each period. The updated interest rate sequence now will be $\eta_r \iota_t + (1-\eta_r) r_t^i$, $\forall t$, with $\eta_r \in (0,1)$.
4. Repeat Steps 2–3 until the interest rate sequence also converges.

As we cannot guarantee the uniqueness of a numerically-constructed competitive equilibrium, we tried many different initial guesses of the wage and interest rate sequences, as well as several values of the relaxation parameters (η_w, η_r). All our competitive equilibria withstood these robustness checks.

B Minimum-Scale Technology

Banerjee and Duflo (2005) point out that models with fixed business start-up costs are more useful in accounting for the observed magnitude of heterogeneity in returns to scale and of misallocation.[24] This point is made in a static environment with generic distortions, and complements the work of Hsieh and Klenow (2007) and Restuccia and Rogerson (2007).

We pursue a related question in this appendix. Given the potential of this minimum-scale technology in accounting for observed misallocation, we ask whether our result on financial frictions and stationary equilibrium remains robust to this technology specification. That is, if financial frictions are the only distortions in an economy with minimum-scale technology, do they have a quantitatively significant effect on the output per capital in the stationary equilibrium?

We give individuals the additional option of operating a more efficient technology with minimum scale of production:

$$f_m(e,k,l) = \Upsilon e^{\nu} \left(\max\{0, k-\underline{k}\}^{\alpha} l^{1-\alpha}\right)^{1-\nu}, \ \Upsilon > 1.$$

[24] A fixed start-up cost or a minimum scale of operation is also considered to be an important feature of technologies associated with the Second Industrial Revolution (Mokyr, 2001). See Mokyr (1990) and especially Braggion (2004) for evidence that financial constraints affected the adoption of new technologies during the Second Industrial Revolution.

For a quantitative assessment of this version with minimum-scale technology, we use a calibration scheme similar to the one in Section 2.1. Note that there are two additional parameters: Υ and $\underline{k}$. We impose $\Upsilon = 1.5$, and choose $\nu = 0.185$, $\zeta = 22.5$ and $\underline{k} = 7.0$ to match the employment share of the top decile, the fraction of entrepreneurs in the population, and the income share of the top five-percentile. The minimum scale of $\underline{k} = 7.0$ is about 38 times the equilibrium wage. All the other parameters are set at their respective values in the benchmark. Note that, other things being equal, the introduction of the minimum-scale technology leads to more concentrated employment distribution. To match the moments of interest, we reduce the dispersion of ability accordingly (i.e. a higher ζ).

In this numerical exercise, the output per capita with $\lambda = 1.25$ is 0.85 times that in the US benchmark with $\lambda = 5.0$. Compared to our earlier result where this output ratio is 0.93 (Table 2), the introduction of a minimum-scale technology does magnify the effect of financial frictions on the output per capita in the stationary equilibrium, but it is still not enough to account for much of cross-country income differences. Again, self-financing turns out to be a good substitute to formal financial markets in the context of the aggregate economy.

Introducing minimum-scale technology has an impact on the transition dynamics as well. In particular, the delaying effect of financial frictions and misallocation of initial capital becomes more pronounced. In the stationary equilibrium, unconstrained high-ability entrepreneurs use the more productive minimum-scale technology. A high-ability type who is born poor now has to operate the less productive technology first and accumulate enough business capital, before utilizing the more productive technology.

References

Acemoglu, D., S. Johnson, and J. A. Robinson (2005): "Institutions as a Fundamental Cause of Long-Run Growth," in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, Amsterdam: Elsevier, vol. 1A, 385–472.

Aghion, P. and P. Bolton (1997): "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies*, 64, 151–172.

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

Albuquerque, R. and H. A. Hopenhayn (2004): "Optimal Lending Contracts and Firm Dynamics," *Review of Economic Studies*, 71, 285–315.

Amaral, P. and E. Quintin (2005): "Financial Intermediation and Economic Development: A Quantitative Assessment," Manuscript, Federal Reserve Bank of Dallas.

Angeletos, G.-M. (2007): "Uninsured Idiosyncratic Investment Risk and Aggregate Saving," *Review of Economic Dynamics*, 10, 1–30.

Atkeson, A. and P. J. Kehoe (1997): "Industry Evolution and Transition: A Neoclassical Benchmark," Working paper, National Bureau of Economic Research.

——— (2005): "Modeling and Measuring Organization Capital," *Journal of Political Economy*, 113, 1026–1053.

Atkeson, A., A. Khan, and L. Ohanian (1996): "Are Data on Industry Evolution and Gross Job Uurnover Relevant for Macroeconomics?" *Carnegie-Rochester Conference Series on Public Policy*, 44, 215–250.

Banerjee, A. V. and E. Duflo (2005): "Growth Theory through the Lens of Development Economics," in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, Amsterdam: Elsevier, vol. 1A, 473–552.

Banerjee, A. V. and A. F. Newman (1993): "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101, 274–298.

Bernanke, B. S. and M. Gertler (1989): "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, 79, 14–31.

Bernanke, B. S., M. Gertler, and S. Gilchrist (1999): "The Financial Accelerator in a Quantitative Business Cycle Framework," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, Amsterdam: North Holland, vol. 1C, 1341–1393.

Blanchard, O. J. (1997): *The Economics of Post-Communist Transition*, Oxford: Clarendon Press.

Braggion, F. (2004): "Credit Market Constraints and Financial Networks in Late Victorian Britain," Manuscript, Tilburg University.

Buera, F. J. (2006): "Persistency of Poverty, Financial Frictions, and Entrepreneurship," Manuscript, Northwestern University.

Cagetti, M. and M. De Nardi (2006): "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, 114, 835–870.

Calvo, G. A. and F. Coricelli (1992): "Stabilizing a Previously Centrally Planned Economy: Poland 1990," *Economic Policy*, 7, 175–226.

Caselli, F. and N. Gennaioli (2005): "Dynastic Management," Manuscript, London School of Economics.

Castañeda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 818–856.

Chen, K., A. İmrohoroğlu, and S. İmrohoroğlu (2006): "The Japanese Saving Rate," *American Economic Review*, 96, 1850–1858.

Christiano, L. J. (1989): "Understanding Japan's Saving Rate: The Reconstruction Hypothesis," *Federal Reserve Bank of Minneapolis Quarterly Review*, 13, 10–25.

Clementi, G. L. and H. A. Hopenhayn (2006): "A Theory of Financing Constraints and Firm Dynamics," *Quarterly Journal of Economics*, 121, 229–265.

Cole, H. L., L. E. Ohanian, A. Riascos, and J. A. Schmitz, Jr. (2005): "Latin America in the Rearview Mirror," *Journal of Monetary Economics*, 52, 69–107.

Deininger, K. and L. Squire (1998): "New Ways of Looking at Old Issues: Inequality and Growth," *Journal of Development Economic*, 57, 259–287.

Duffie, D., J. Geanakoplos, A. Mas-Colell, and A. McLennan (1994): "Stationary Markov Equilibria," *Econometrica*, 62, 745–781.

Erosa, A. (2001): "Financial Intermediation and Occupational Choice in Development," *Review of Economic Dynamics*, 4, 303–334.

Evans, D. S. and B. Jovanovic (1989): "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy*, 97, 808–827.

Galor, O. and J. Zeira (1993): "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60, 35–52.

Giné, X. and R. M. Townsend (2004): "Evaluation of Financial Liberalization: A General Equilibrium Model with Constrained Occupation Choice," *Journal of Development Economics*, 74, 269–307.

Greenwood, J. and B. Jovanovic (1990): "Financial Development, Growth, and the Distribution of Income," *Journal of Political Economy*, 98, 1076–1107.

Guner, N., V. Gustavo, and X. Yi (2006): "Macroeconomic Implications of Size-Dependent Policies," Manuscript, Pennsylvania State University.

Hall, R. E. and C. I. Jones (1999): "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114, 83–116.

Hsieh, C.-T. (2002): "What Explains the Industrial Revolution in East Asia? Evidence from the Factor Markets," *American Economic Review*, 92, 502–526.

Hsieh, C.-T. and P. Klenow (2007): "Misallocation and Manufacturing TFP in China and India," Manuscript, Stanford University.

Huggett, M. (1997): "The One-Sector Growth Model with Idiosyncratic Shocks: Steady States and Dynamics," *Journal of Monetary Economics*, 39, 385–403.

Jeong, H. and R. M. Townsend (2005): "Growth and Inequality: Model Evaluation Based on an Estimation-Calibration Strategy," Manuscript, University of Southern California.

——— (2007): "Sources of TFP Growth: Occupational Choice and Financial Deepening," *Economic Theory*, 32, 197–221.

King, R. G. and S. T. Rebelo (1993): "Transitional Dynamics and Economic Growth in the Neoclassical Model," *American Economic Review*, 83, 908–931.

Kiyotaki, N. and J. Moore (1997): "Credit Cycles," *Journal of Political Economy*, 105, 211–248.

Klenow, P. J. and A. Rodríguez-Clare (1997): "The Neoclassical Revival in Growth Economics: Has it Gone Too Far?" in *Macroeconomics Annual 1997*, ed. by B. S. Bernanke and J. J. Rotemberg, Cambridge: MIT Press, 73–102.

Krusell, P. and A. A. Smith, Jr. (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

Lloyd-Ellis, H. and D. Bernhardt (2000): "Enterprise, Inequality and Economic Development," *Review of Economic Studies*, 67, 147–168.

Lucas, Jr., R. E. (1978): "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9, 508–523.

Matsuyama, K. (2007): "Aggregate Implications of Credit Market Imperfections," in *NBER Macroeconomics Annual*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford, Cambridge, MA: MIT Press, vol. 22.

Miao, J. (2006): "Competitive Equilibria of Economies with a Continuum of Consumers and Aggregate Shocks," *Journal of Economic Theory*, 128, 274–298.

Mokyr, J. (1990): *The Lever of Riches: Technological Creativity and Economic Progress*, New York and London: Oxford University Press.

——— (2001): "The Rise and Fall of the Factory System: Technology, firms, and households since the Industrial Revolution," *Carnegie-Rochester Conference Series on Public Policy*, 55, 1–45.

Morley, S. A., R. Machado, and S. Pettinato (1999): "Indexes of Structural Reform in Latin America," Serie de reformas económicas, United Nations Economic Commission for Latin America and the Caribbean, Santiago, Chile.

Mukand, S. W. and D. Rodrik (2005): "In Search of the Holy Grail: Policy Convergence, Experimentation, and Economic Performance," *American Economic Review*, 95, 374–383.

Parente, S. L. and E. C. Prescott (2000): *Barriers to Riches*, Cambridge: MIT Press.

Piketty, T. (1997): "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing," *Review of Economic Studies*, 64, 173–189.

Quadrini, V. (2000): "Entrepreneurship, Saving, and Social Mobility," *Review of Economic Dynamics*, 3, 1–40.

Restuccia, D. and R. Rogerson (2007): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," Working Paper 13018, National Bureau of Economic Research.

Roland, G. (2000): *Transition and Economics: Politics, Markets, and Firms*, Cambridge: MIT Press.

Sokoloff, K. L. and S. L. Engerman (2000): "History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives*, 17, 217–232.

Solon, G. (1999): "Intergenerational Mobility in the Labor market," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter and D. Card, Amsterdam: North Holland, vol. 3, 1761–1800.

Young, A. (1995): "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience," *Quarterly Journal of Economics*, 110, 641–680.