China’s Great Convergence and Beyond

Kjetil Storesletten1 and Fabrizio Zilibotti2

1Department of Economics, University of Oslo, Oslo NO-0317, Norway; email: kjetil.storesletten@econ.uio.no
2Department of Economics, University of Zurich, Zurich CH-8006, Switzerland

Keywords
economic reforms, foreign imbalances, innovation, middle-income trap, pension reform, savings

Abstract
A recent wave of economic research has studied the transformation of China from a poor country in the 1970s to a middle-income economy today. Based on this literature, we discuss the factors driving China’s development process. We provide a historical account of China’s rise, fall, and resurgence. We then discuss the stylized facts associated with China’s growth process and review a comprehensive theory of its economic transition. Finally, we discuss China’s future. In particular, we review some recent studies about technological and politico-economic factors that may foster or hinder its future economic performance.
1. INTRODUCTION

China is today one of the world’s most powerful nations. China’s population of 1.36 billion exceeds that of high-income nations altogether, and the country is presently the world’s second largest economy and the largest exporter.\(^1\) However, until the late 1970s, China was a very poor and closed country, with an income per capita of just 4% of the US level. The process of economic reform, which started in the 1980s and accelerated in the 1990s, catapulted China onto a trajectory of stellar growth. Over the first decade of the twenty-first century, China’s (purchasing power parity adjusted) GDP per capita grew at an annual rate of approximately 9%. The resulting reduction in poverty is equally impressive: The fraction of the Chinese population living in extreme poverty plummeted from 84% in 1981 to 13% in 2008.\(^2\) Such a rapid improvement in living conditions for such a large share of the world population is unprecedented in history. Although China is much richer today than it was 30 years ago, its level of economic development is still low, with an income per capita of approximately 20% of the US GDP per capita.

As the Middle Kingdom has gained in economic importance, it has attracted increasing attention from the economic profession. Understanding the driving forces behind the most significant growth miracle in history is indeed a first-order issue. Yet China remains largely understudied. The number of studies on China published in general interest journals is still very small relative to the number of those focusing on the United States (or other industrialized countries). This is largely a result of conservatism and inertia in the profession. However, this will change in the coming years. In our view, the interest in understanding China’s great economic transformation will remain strong and be more resilient than interest in other current issues such as the Great Recession.

In this review, we address three related questions: (a) Why was China so poor in the 1970s? (b) Why was growth so rapid after 1979? (c) What factors can sustain and what can jeopardize the future growth of China? We do not provide comprehensive answers to these big questions. Instead, we highlight some selected issues we regard as salient and fruitful for further economic research.

The article is structured as follows: Section 2 reviews China’s historical development during the second millennium. Section 3 presents the stylized facts on the transition of China toward a market economy and lays out a theoretical framework that captures these observations. Section 4 discusses issues related to the future economic development of China. Section 5 concludes.

2. HISTORICAL BACKGROUND

2.1. From World Technological Leadership to Poverty

Between 1000 and 1500 AD, China was the most technologically advanced region worldwide. Prosperity stretched from the Song period (960–1279) to the commercial development under the Ming dynasty (1368–1644), when sea explorations led Chinese traders all the way to the coasts of Africa (Levathes 1994, pp. 1405–33). During this period, China introduced many important inventions that would become known in Europe centuries later. The so-called four great inventions—printing, gunpowder, paper making, and the compass—were only some of the major innovations introduced during the Song period or earlier (Needham 1981, Shaffer 1986, Lin 1995). Under that same dynasty, the central administration started issuing paper money, well ahead of Europe. As long-distance trade with Europe developed in the sixteenth century, China exported technology-intensive

\(^1\)The United States, Western and Central Europe, Japan, and other Western offshoots altogether have a population of approximately 1 billion.

\(^2\)This is using the World Bank’s definition of poverty (i.e., those living with less than $1.25 daily).
goods in exchange for silver and primary commodities (Wong 1998). The demographic evolution kept pace with the general prosperity: By 1100, the Chinese population rose to over 100 million, reaching 160 million at the time of the Ming dynasty (Fairbank & Goldman 2006, p. 128).

After the Ming splendor, power was seized by the Qing dynasty (1644–1912), native of Manchuria. It took the new rulers about four decades to conquer the whole country and to crush the Ming resistance. The conflict plunged the country into a severe economic downturn. Yet China recovered, and by the end of the seventeenth century, the economy was flourishing again. According to Pomeranz (2000, 2001), the living standards in the Yangtze Delta region—the richest part of the country—were, in the second half of the eighteenth century, comparable to those of the most advanced European regions.3

The British Industrial Revolution was the landmark of the great divergence. Hostile to Western influence, the imperial government imposed heavy barriers to the commercial relationships with Europe.4 The ensuing conflict with the Western colonial powers led to a sequence of wars ending in military defeats for China (e.g., the Opium Wars of 1839–1842 and 1856–1860). In turn, these severely undermined the legitimacy of the imperial government. Unrest erupted, most notably the Taiping Rebellion, an outright civil war, which brought the Manchurian rulers to confront surging Han nationalism between 1851 and 1864 (Platt 2012). The revolt ended in bloodshed. Between 20 and 30 million people are estimated to have died as the army repressed the revolt with the help of the French and English armies.

China benefitted, as did Britain, from important technological improvements in agriculture preceding the Industrial Revolution. The control of river floods caused a surge in food production, which in turn induced a demographic boom: Between 1680 and 1820, the population tripled. However, unlike in the West, progress in agriculture did not pave the way to industrialization and urbanization. In Britain, these improvements preluded the breakdown of the Malthusian equilibrium. Fertility started to decline and income per capita grew. China, in contrast, remained a rural country, with a stagnating income per capita.5

This great divergence accelerated when Chinese political institutions collapsed in the early twentieth century. A revolution in 1911 led to the proclamation of the republic under the presidency of Sun Yat-sen. However, the new state was weak and precipitated into a period of wars and anarchy.

In 1949, after the end of the Second Sino-Japanese War and the communist uprising, the People’s Republic of China (PRC) was founded under the leadership of Mao Zedong, chairman of the Chinese Communist Party (CCP). The PRC was a vastly impoverished country, dominated by traditional subsistence activities. From 1951 and onward, industry and agriculture were collectivized. Dissatisfied with the slow speed of progress, Mao launched in 1956 the Hundred Flowers Campaign, inviting intellectuals and ordinary people to voice their open criticism of the Party’s policies and bureaucracy. This window for open debate was soon closed and replaced by the call for a Great Leap Forward, an ambitious (and improvised) plan intended to turn the PRC into a modern industrial collectivized country. Its implementation contributed to the ensuing famine that killed approximately 30 million people (Meng et al. 2010). After an ephemeral reform-oriented

3Other economic historians challenge this finding, arguing that living standards in the richest Chinese city, Beijing, were already below those of London and Amsterdam, although were comparable to cities such as Leipizig and Milan in the eighteenth century (Allen et al. 2011).

4An example is the restriction of foreign trade to the port of Canton, subject to high duties (see Krieger et al. 1990 and Wong 2004 for a different interpretation of these restrictions).

5For different interpretations of the divergence between China and Europe, readers are referred to, for example, Pomeranz (2000) and Voigtlander & Voth (2013).
stage under the aegis of Liu Shaoqi and Deng Xiaoping (1962–1964), a new wave of radicalism erupted with the start of the Cultural Revolution, which was supposed to cleanse the society of capitalism and traditional Chinese values. During this tormented period, Liu Shaoqi was jailed and died in prison, and Deng Xiaoping fell into disgrace. After Mao’s death in 1976 and the liquidation of the Gang of Four—a group of leftist party officials, including Mao’s wife—Deng Xiaoping became the de facto leader of the CCP. He quickly repudiated the Cultural Revolution and in 1978 launched a program of pragmatic economic reforms whose primary goal was to increase the persistent low productivity in agriculture.6

2.2. The 1980s: Experimenting with Economic Reforms

After 30 years of central planning, the PRC gradually adopted a set of market-oriented reforms. Land collectivization was replaced by the principle of household responsibility in agriculture, and the role of local governments was enhanced by the creation of township and village enterprises. The government experimented with a new industrial policy granting a special status to a few selected Special Economic Zones (SEZs). The decade marked the start of a rapid structural transformation from agriculture to industry and services: Between 1978 and 2003, the employment share of agriculture fell from over 70% to less than 50% (Dekle & Vandenbroucke 2012, Cao & Birchenall 2013).

SEZs were especially important. They opened a hitherto isolated economy to foreign investments and the associated flow of technical knowledge. Initially, four SEZs were established: Shenzhen, Zhuhai, Shantou in the Guangdong Province, and Xiamen in the Fujian Province. The success of the experiment led to a progressive expansion of Chinese industrial policy: In 1984, 14 cities on the east coast and later two provinces and three delta areas became SEZs. Then, in 1992, 1998, and 2005, the SEZ status was extended to inland cities: first capitals, then median cities.

In a recent empirical study, Alder et al. (2013) assess the effects of SEZs on economic development. The establishment of SEZs introduced well-defined changes to the legal framework, which were staggered over time and space.7 The authors exploit this variation in a panel of 276 Chinese cities at the prefecture level over the period 1988–2010 to estimate, by a difference-in-difference method, the effect of the policy treatment.8 The control group includes cities in the same province and year that have not (or not yet) received the SEZ status. The authors find that introducing an SEZ increases the city’s GDP per capita by 20% in the long run.9 Figure 1 displays the average treatment effect of SEZs on GDP per capita before and after their establishment.

One might be concerned that the government’s choice of SEZ location is not random. For example, the central government may have selected cities based on some prior knowledge (e.g., proximity to ports, and thus potential for success). Indeed, in the 1980s, SEZs were selected on the coast and close to Hong Kong and Taiwan. To address this selection issue, Alder et al. (2013)

---

6The pragmatic attitude of Deng Xiaoping is captured by his quote, “I don’t care if it’s a white cat or a black cat. It’s a good cat so long as it catches mice.” He is said to have made this statement during the 1961 Great Famine, speaking in favor of partial decollectivization of land so as to raise agricultural production (Li 1994, p. 376).

7SEZs received preferential treatment in terms of tax deduction, custom duty deduction, reduced land-use price, and flexibility in labor and financial contracts.

8Earlier studies documenting cross-sectional growth differences between cities hosting SEZs and other cities include Wei (1993), Demurger et al. (2002), and Jones et al. (2003). J. Wang (2013) shows that SEZs increased the FDI flow.

9Alder et al. (2013) estimate the following regression: $y_{ipt} = \phi_i + \gamma_{t,p} + \sum_{n=1}^{T} \alpha_n I_{ipt}(t - \text{Reformyear}_t = n) + X_{ipt} \beta + \epsilon_{ipt}$, where $t - \text{Reformyear}_t$ measures leads and lags of the reform year in city $i$. The regression controls for city fixed effects $\phi_i$ (to filter out time-invariant heterogeneity), province time fixed effects $\gamma_{t,p}$ (to control for time-varying heterogeneity at the provincial level), and other city characteristics $X_{ipt}$.
restrict the analysis to a subsample of treated cities located in inland provinces where the assignment was based on administrative criteria (e.g., being a provincial capital). The effects for this subsample are similar to the rest. Moreover, Figure 1 shows that there are no differential trends prior to reforms.

The evidence in Alder et al. (2013) suggests that the industrial policy was effective in promoting economic development. The quantitative effects of SEZs on city development are sizeable but not huge relative to the high total growth rates of this period, suggesting that the effects of the policies might have spread relatively quickly to neighboring regions. Even beyond its direct economic effects, the SEZ experiment made a significant contribution. It strengthened the confidence of the political leadership that opening the Chinese economy to markets and trade was essential to achieve industrialization and economic development. Two decades after the first SEZs were established, the development strategy of the political leadership still echoes the positive experience of the SEZs: On September 29, 2013, the government launched the Shanghai Free Trade Zone, with an emphasis on financial liberalization. We return to this recent development below.

In summary, the 1980s is an important decade of experimentation in economic reform. Development took off as productivity in agriculture soared and domestic and foreign investments poured into the SEZs. Nevertheless, most of the country remained subject to a centralized planning system. By the end of the decade, growth slowed down and popular discontent grew as widespread corruption within the political elite was exposed. Street demonstrations after the death of the reformist leader Hu Yaobang in 1989 were followed by student protests, violently repressed, in Tiananmen Square. Divisions fermented within the CCP as its conservative faction challenged the whole process of economic reform.

### 2.3. The 1990s: China Becomes a Market Economy

The early 1990s was marked by great uncertainty. Eventually, the proreform faction won, and the reform process resumed. The turning point was Deng Xiaoping’s Southern Tour of 1992, a series of informal speeches laying out the Party’s new course and gathering support for deeper economic reforms that would transform China irreversibly (Zhao 1993). In 1990, the Shanghai Stock
Exchange reopened. State-owned enterprises (SOEs) were subjected to market competition, with inefficient and unprofitable entities forced to either restructure or shut down. Many SOEs were privatized or formed joint ventures with foreign firms. Wholly private enterprises owned by Chinese entrepreneurs were created and received the official blessing of the CCP in 1997. China embraced a process of export-led growth, culminating in accession to the World Trade Organization in 2001.

It is often argued that the growth in the 1990s was largely driven by investment. Indeed, investment rates were sustained at a very high level, well over 30%. However, the economic transformation went far beyond mere capital accumulation: It brought about major shifts in the sectoral composition of output and urbanization and a growing importance of markets, technological change, and entrepreneurial skills. Without these changes, it would have been impossible for the rate of return on investments to remain so high over such a long period. \(^{10}\) If part of this success can be explained by the adoption of new technologies from existing firms, reallocation has been a key driver of the growth process. For instance, Brandt et al. (2012b) document that up to two-thirds of the aggregate total factor productivity (TFP) growth in manufacturing resulted from selection (i.e., high-productivity firms entering and low-productivity firms exiting the market).

The lion’s share of this reallocation process is the exit of low-productivity SOEs and their replacement with new domestic private enterprises (DPEs). Figure 2 shows the private employment share in manufacturing, mining, and construction, including both DPEs and foreign-owned enterprises. In 1994, private enterprises accounted for approximately 10% of total employment. By 2007, their share exceeded 50%. SOEs are on average less productive than DPEs. Song et al. (2011) report a difference of 9 percentage points in profitability between DPEs and SOEs, measured by the ratio between total profits (“operation profits plus subsidies plus investment returns”) and the value of fixed assets net of depreciation. Similar results emerge from TFP accounting studies, such as Brandt et al. (2008), Brandt & Zhu (2010), and Brandt et al. (2012a). Hsieh & Klenow (2009) estimate a measure of the revenue-TFP gap between DPEs and SOEs to be 42%. In summary, the process of reallocation is related intimately to the progressive privatization of the Chinese economy.

Another salient feature of the Chinese transition is the moderate wage growth. The average real annual growth of wages in the urban manufacturing sector was 7.6% from 1992 to 2007, whereas the average growth rate in the urban real GDP per capita during the same period was above 10% (Ge & Yang 2014). Moreover, part of the measured wage growth stems from a composition effect as the share of educated workers has risen. There is also evidence of a falling labor share of aggregate output, from 50% in 1992 to 41% in 2005 (Bai et al. 2006, table 1). The relatively low wage growth, combined with a thriving class of entrepreneurs, has contributed to the rising inequality in China.

Finally, important aspects of the Chinese growth experience are pervasive credit and financial market imperfections, as documented, for example, by Allen et al. (2005) and Song & Wu (2011). A symptom of these imperfections is the large gap between high corporate returns and very low returns on savings: The average real rate of return on bank deposits, the main financial investment of Chinese households, was close to zero. Firms are credit constrained, with private firms more so than SOEs. This discrimination is reflected in the fact that SOEs finance a substantially larger share of new investments through external channels (see Song et al. 2011). Moreover, Dollar & Wei

\(^{10}\) China’s rate of return on capital has remained well above 20%, higher than in most industrialized and developing economies. Bai et al. (2006, figure 11) show that the rate of return on investment in the manufacturing sector has increased since the early 1990s.
(2007) document survey evidence that private firms must often resort to retained earnings and family or friends to finance investments. Long & Zhang (2011) provide evidence that private firms benefit more than do state-owned firms from industrial clusters, as these facilitate interfirm trade credit, easing their need for external financing. The difficulty for DPEs to finance investments is also reflected in the choice of production techniques. In 2006, the average capital-output ratio in SOEs was 1.75 versus 0.67 in DPEs, while the capital per worker was almost five times larger in SOEs than in DPEs. This difference has both an intensive and an extensive margin. Song et al. (2011) document that SOEs are more capital intensive within three-digit manufacturing industries. Moreover, DPEs specialize in labor-intensive industries, whereas SOEs still dominate capital-intensive industries.

3. GROWING LIKE CHINA

In this section, we present a model that captures the salient features of China’s transition into a market economy, outlined above. The model is based on Song et al. (2011).

3.1. A Theory of Economic Transition

Both the neoclassical and the endogenous growth literatures developed in the 1990s focus on a representative firm and study its incentives to accumulate capital and technical knowledge (see, e.g., Aghion & Howitt 1992). In the analysis of developing economies, growth theory focuses on the imitation and adoption of more advanced technologies already in use in other countries (Acemoglu et al. 2006). The lag in technology adoption determines the TFP gap across countries.
A recent literature argues, however, that low aggregate TFP, especially in developing countries, can arise from firm-level misallocation (see, e.g., Restuccia & Rogerson 2008, Gancia & Zilibotti 2009, Hsieh & Klenow 2009). Although some firms have high productivity, they fail to attract large shares of productive resources, owing to financial frictions and other wedges. These wedges keep inefficient firms alive, reducing average productivity. Hsieh & Klenow (2009) show that reallocation within the manufacturing sector has been an important driver of productivity growth in China, accounting for an annual 1.4–percentage point increase in aggregate TFP during 1998–2005.

Song et al. (2011) provide a framework for analyzing the growth effects of a large initial misallocation that is removed gradually over time (as in Lewis 1954). The building blocks of the theory are firm heterogeneity in productivity and financial market imperfections. There are two types of firms: Entrepreneurial E-firms have a higher TFP and are operated by agents with entrepreneurial skills who are financially constrained. Financially integrated F-firms have a lower TFP but have good access to credit markets. In particular, the latter can borrow at the world interest rate, assumed to be constant. In the Chinese case, E-firms and F-firms capture DPEs and SOEs, respectively. In addition to financial frictions, we introduce a labor wedge implying that labor costs may differ across the two types of firms. Before economic reforms, such a wedge is large, implying that there is no employment in E-firms. The economic reforms reduce this wedge, triggering the onset of privatization. Absent credit market frictions, F-firms would be instantaneously crowded out by the more productive E-firms. However, F-firms temporarily survive, owing to their better access to credit markets. Over time, the self-financed E-firms outgrow the F-firms, progressively reducing their employment share. During this transition, the average rate of return to capital and TFP increase owing to reallocation.

More formally, the technologies operated by F-firms and E-firms are described by the following production functions:

$$y_{Ft} = k_{Ft}^{\eta} (A_t n_{Ft})^{1-a} \quad \text{and} \quad y_{Et} = k_{Et}^{\eta} (\chi A_t n_{Et})^{1-a},$$

where $k$ and $n$ denote the capital and labor input, respectively. $\chi > 1$ reflects the better governance of E-firms.$^{11}$ $A_t$ is a TFP parameter that evolves according to an exogenous law of motion: $A_{t+1} = (1 + z)A_t$.

F-firms can borrow from banks at the gross rate $R^i$. Profit maximization implies that $R^i$ equals the marginal product of capital in F-firms and that wages equal the marginal product of labor,

$$w_t = (1 - a) \left( \frac{a}{R^i} \right)^{\frac{1}{\alpha}} A_t. \quad (1)$$

E-firms must hire a manager to run the firm. The manager’s compensation, $m_t$, is subject to a standard no-stealing incentive-compatibility constraint requiring that, for some $\psi \in (0, 1)$, $m_t \geq \psi y_{Et}$. Subject to some parameter restriction, this constraint is binding in equilibrium. Thus, the value of a firm endowed with capital $k_{Et}$ is the solution to the following program:

$$\Xi_t(k_{Et}) = \max_{n_t} \left\{ (1 - \psi) \left[ (k_{Et})^{\alpha} (\chi A_t n_{Et})^{1-a} - m_t - (1 + \omega)w_t n_{Et} \right] \right\}, \quad (2)$$

where $\omega$ is a labor market wedge faced by private firms. $\omega > 0$ may reflect subsidies to F-firms or costly regulations that affect only E-firms (in which the complete ban on domestic private firms

---

11Readers are referred to Song et al. (2011) for microfoundations. The weak corporate governance of SOEs is well documented (see Chang & Wong 2004, Liu & Otsuka 2004).
prior to the 1990s would correspond to \( \omega \to \infty \). In contrast, a negative wedge can arise if E-firms can avoid some costs (e.g., by evading payroll taxes or contributions to the pension system).

Taking the first-order condition with respect to \( n_E \) and substituting in the equilibrium wage given by Equation 1 implies that employment in E-firms is linear in their capital stock,

\[
n_E t = \left( \frac{1 - \psi}{1 + \omega} \right)^{\frac{1}{\alpha}} \left( \frac{c}{R^I} \right)^{\frac{1}{\alpha}} \frac{k_{Et}}{\chi A_t}.
\] (3)

This in turn implies that the rate of return to capital invested in E-firms is constant, \( \rho_E = (1 - \psi)^{1/\alpha} \left( \chi/(1 + \omega)^{(1 - \alpha)/\alpha} \right) R^I \), and that the value of the firm is proportional to the capital stock, \( \Xi_t(k_{Et}) = \rho_E k_{Et} \). We note that the firm-level results also hold at the aggregate level.

Banks collect savings from households and invest in domestic capital and foreign bonds, yielding a gross return \( R \). We assume a stark financial friction: Banks do not lend to E-firms (in Song et al. 2011, E-firms can borrow to finance investments, but only up to a limit). An arbitrage condition implies that the rate of return on domestic investments equals the rate of return on foreign bonds. Moreover, assuming for simplicity a competitive banking sector yields \( R_d = R_l = R \), where \( R_d \) is the return on deposits.

The household sector is populated by overlapping generations of two-period-lived agents who work when young and live off savings when old. Preferences are parameterized by a standard logarithmic utility function, \( U_t = \log (c_{1t}) + \beta \log (c_{2t+1}) \), where \( \beta \) is the discount factor. Each agent saves a fraction \( \beta/(1 + \beta) \) of his or her first-period income.

Agents have heterogeneous skills. Each cohort consists of a measure \( N_t \) of agents with no entrepreneurial skills (workers) and a measure \( m_t \) of agents with entrepreneurial skills (entrepreneurs). Skilled agents can be hired as managers when young and invest in E-firm capital when old. The population of workers grows at the exogenous rate \( n \), where \( n \) captures demographic trends, including migration from rural to urban areas.

Consider the young households’ portfolio choice. Workers face no interesting choice and deposit their savings with banks, earning the interest rate \( R_d \). Young managers can invest their savings either in bank deposits or in E-firm capital. Because there is no risk, they invest all their savings in the firm if and only if \( \rho_E > R_d \). This occurs if and only if

\[
\frac{\chi}{1 + \omega} > \left( \frac{1}{1 - \psi} \right)^{\frac{1}{\alpha}}.
\] (4)

This is a necessary condition for economic transition to take off. Conditional on \( \chi \) and \( \psi \), the transition requires that the wedge \( \omega \) be sufficiently low.

Suppose, first, that the condition in Equation 4 fails to hold. An extreme example is China before the start of economic reforms, when private employment contracts are forbidden. Then, the economy is in a steady state in which all workers are employed in F-firms and productivity grows at the exogenous rate \( z \). Next, assume that \( \omega \) falls, owing to economic reforms, so that the condition in Equation 4 is satisfied. The policy change triggers a transition in which skilled agents set up E-firms and hire workers and managers, and capital starts accumulating in the E-firm sector.

\[12\text{For the transition to take off, one more condition must be met, namely that the entrepreneurs must have enough initial capital that the wage of managers exceeds that of workers. In our model, the number of firms is indeterminate. We can therefore assume that skilled agents can (if necessary) pool their savings so as to form a firm that is sufficiently large that } m > w. \text{ Alternatively, the initial capital can come from the privatization of some SOEs.}\]
The transitional equilibrium is analyzed more formally by Song et al. (2011), who show that, given \( K_{E_t} \) and \( A_p \), the equilibrium dynamics of total capital and employment among E-firms during transition are given by 
\[
\frac{K_{E_{t+1}}}{K_{E_t}} = 1 + \gamma \quad \text{and} \quad \frac{N_{E_{t+1}}}{N_{E_t}} = \frac{(1 + \gamma)/(1 + z)}{(1 + \gamma)/(1 + z)},
\]
respectively, where
\[
1 + \gamma = \frac{\beta \psi}{1 + \beta} \frac{\rho_E}{\alpha}.
\]

For E-firms’ share to grow over the transition, it is necessary that 
\[
\frac{(1 + \gamma)/(1 + z)}{(1 + \gamma)/(1 + z)} > 1 + \nu.
\]
The speed of transition is increasing in \( \beta \) and \( \chi \) and is decreasing in \( \omega \). Intuitively, a high propensity to save for the entrepreneurs yields high investments in E-firms, speeding up the transition. So does a higher TFP in E-firms. The wedge \( \omega \) has the opposite effect.

Figure 3 illustrates the transitional dynamics of the economy in Song et al. (2011). During the transition, the E-firms’ employment share grows. Once all workers are employed in E-firms, the transition ends (period \( T \) in the figure). E-firms have a higher rate of return and less capital per worker than do F-firms. The within-firm rates of return on capital remain constant, but the average rate of return on capital increases due to a composition effect. The growth of GDP per
worker increases during the transition, due to the reallocation of capital and labor toward more productive firms. The wage rate grows at the rate $z$, so the wage per effective unit of labor remains constant.

These observations are consistent with the stylized facts described in Section 2. In particular, the employment share of DPEs has been increasing (see Figure 2), the rate of return on capital in manufacturing has been increasing, the wage growth has been significantly lower than output growth, and DPEs have a higher return and less capital per worker than SOEs. In addition, output growth increases during the transition, consistent with the empirical pattern for China, and then eventually declines. Interestingly, there are current signs of a growth slowdown, accompanied by a reduction in the reallocation of workers from SOEs to DPEs (see Figure 2).

An implication of the theory is that liberalization reforms that relax asymmetric financial constraints can foster the growth of private firms and speed up transition. Fan & Kalemli-Ozcan (2014) provide empirical support for this prediction. They study the effects of financial reforms on firm-level investments in a panel of Asian emerging economies, including China, using a difference-in-difference approach. They find that financial liberalization reduces the savings and increases the investments of private firms relative to those of state-owned firms. This is consistent with the premise of the theory above that private firms are subject to tighter financial constraints. Reforms relaxing such constraints reduce the need for private firms to retain earnings to fund their investments and increase their investment potential.\footnote{Fan & Kalemli-Ozcan’s (2014) findings suggest that asymmetric financial frictions are pervasive in East Asian economies beyond China. Song et al. (2011) document that, in this respect, the development of China is similar to the experiences of Korea and Taiwan.}

Finally, the model predicts increasing inequality and a declining labor share. In the prereform economy, the labor share of output equals $1 - \alpha$, whereas at the end of the transition, it is equal to $(1 - \psi)(1 - \alpha)$. Although wages grow at the rate $z$, significantly below the average growth rate in the economy, managerial compensations grow at the rate $\gamma > z$ during the transition: A new middle class of entrepreneurs thrives.

In the theory outlined above, the transition is ignited by the change in the employment wedge in the private sector, $\omega$. What were these policy changes that triggered growth? The success of the reform process cannot be attributed to one policy measure alone. As discussed in Section 2, the industrial policy based on SEZs may have been important. However, the size of Alder et al.’s (2013) estimated effects suggests that this cannot have been the only factor at work. A complementary theory is that a change in the system of incentives within the ruling CCP created conditions conducive for growth. Jia (2012) provides evidence that in the 1990s, economic growth became the single most important criterion for political careers within the CCP. We return to her study in Section 4.

### 3.2. Zhuada Fangxiao

The model outlined above captures salient aspects of economic transition in China. However, the stark prediction that SOEs disappear altogether is not fully borne out in the data (see Figure 2). Indeed, the strategy laid out by the Ninth Five-Year Plan in 1997 was not to aim for a privatization of all sectors of the economy, but rather to retain a strong state presence in some key sectors and to rationalize the system of SOEs. The slogan was “grab the large and release the small firm” (zhuada fangxiao). In the data, we see that the role of SOEs has declined substantially in many sectors, especially labor-intensive ones, but remains dominant in other sectors, especially capital-intensive industries (e.g., electrical and heating power or transport equipment). In these industries, the
government promoted the merger and restructuring of SOEs into large transregional groups. Moreover, there is evidence of rising profits (and wages) in surviving SOEs. Hsieh & Song (2013) show that the TFP growth of surviving SOEs has been high and even greater than that of DPEs (although a large productivity gap in favor of DPEs persists).

Song et al. (2011) provide two complementary explanations for these developments. First, a multisector extension of the model in which some sectors are more capital intensive than others shows that market forces drive SOEs out of business more quickly in labor-intensive industries, whereas SOEs are more resilient in capital-intensive industries. Intuitively, the SOEs’ access to external financing gives them a comparative advantage in capital-intensive activities. Second, if SOEs are granted monopoly power in specific industries, they will not only remain dominant in those industries, but will also benefit from the increased efficiency of the rest of the manufacturing industries when these are liberalized.

We illustrate the monopoly story here. Consider the following extension of the model above. Assume the final good, $Y_f$, to be a constant elasticity of substitution aggregate of two intermediate goods:

$$Y_f = \left( \varphi \left( \frac{Y_c}{\pi} \right)^{\frac{1}{\varphi + 1}} + \left( \frac{Y_m}{\pi} \right)^{\frac{1}{\varphi + 1}} \right)^{\frac{\varphi + 1}{\varphi}}. \quad (6)$$

Initially, both $Y_c$ and $Y_m$ are produced by F-firms. Then, economic reforms liberalize production in the $Y_c$ sector (which becomes competitive), which can be interpreted as a reduction in $\omega$ in this sector. However, no reform takes place in the $Y_m$ sector (i.e., SOEs retain monopoly in this sector). Over time, DPEs take over the competitive sector, according to the mechanism of the general model outlined above. Song et al. (2011, section IV, proposition 2) show that the optimal markup charged by the monopolist SOEs (and, hence, their profits) is increasing in the share of E-firms in the competitive industry. Intuitively, as the productivity of the competitive sector increases during the transition, so does the demand for the monopolized good, which increases the profits of the monopolist.

Interestingly, the model predicts a nonmonotonic behavior in SOEs in terms of profitability. Before the reform, both sectors are monopolized by the SOEs. Then, as sector $c$ becomes exposed to competition from private firms, the SOEs in the competitive sector stop making profits. In this stage, the average performance of SOEs deteriorates. Eventually, as DPEs crowd out the SOEs from the competitive sector, the average performance of surviving SOEs improves again. The reason is that the vast majority of the surviving SOEs are now in the monopolized sector, and in addition, as discussed above, profits in this sector are large because of the high efficiency attained by the other sector.

A complementary theory along similar lines was developed recently by Li et al. (2012). They assume a vertical structure in which SOEs monopolize upstream industries, whereas the downstream industries are competitive. Similar to the mechanism in Song et al. (2011) discussed above, the rents of the upstream SOEs increase because of the liberalization of the downstream sectors. The authors interpret the model economy as one of “state capitalism.”

Although Song et al. (2011) assume that labor markets are competitive, so all firms pay equal wages, there is evidence that SOEs pay on average higher wages for workers of a given qualification. Moreover, the gap has increased over time, according to Ge & Yang (2013). It would be simple to extend the model to allow for ex post rent sharing (e.g., through Nash bargaining) between workers and firms in the monopolized sector. Such a model predicts that as the profits of surviving SOEs grow, so do the wages paid by these firms. Another reason why SOEs may decide to grant part of the rents to their workers may be political motives (i.e., the desire to create a base of support for the government elite). We return to this argument below.
Finally, there is evidence that surviving SOEs have managed to become more efficient. As mentioned above, Hsieh & Song (2013) document that TFP grew faster in SOEs than in private firms during 1998–2007 (see also Liu & Cao 2011). This is partly the result of improvements in the management of surviving SOEs and partly a result of positive selection: The least productive (typically small) SOEs have exited the market or have been forced to merge or restructure, whereas the most productive (typically large) SOEs have survived. Improvements in SOEs have been quantitatively sizeable. Hsieh & Song (2013) perform a counterfactual experiment in which the TFP growth of all incumbent SOEs is shut down, internal distortions (in particular, labor distortions) are set equal to the 1997 levels for all incumbent SOEs, and all SOEs that have exited are resurrected. The result is that total industrial output would be 42% lower.

3.3. The Foreign Surplus

A feature of the Chinese experience during the past two decades is that high growth has been accompanied by the accumulation of a large foreign surplus. Its foreign reserves swelled from $21 billion in 1992 (5% of its GDP) to $3,500 billion in June 2013 (over 40% of its GDP) (see Figure 4).

The recent literature has noted the regularity that many emerging economies have a large foreign surplus in spite of higher return on investments in the rest of the world. Gourinchas & Jeanne (2014) show that on average, countries with fast TFP growth have large trade surpluses and capital outflows, whereas countries with low TFP growth have trade deficits and capital inflows. They label this finding the “allocation puzzle,” as it runs against the predictions of the standard neoclassical theory.

![Figure 4](https://www.annualreviews.org/doi/abs/10.1146/annurev-economic-020113-155551)

**Figure 4**
China’s foreign reserves (solid line) and the domestic bank deposits minus domestic loans (dotted line), both measured as a share of GDP. Data taken from *China Statistical Yearbook*, various issues.
Although China is by no means unique, the foreign surplus of China has drawn far greater public attention than the surplus of other countries, owing to its large weight in the world economy. A concern is that cheap Chinese exports harm domestic firms and cause job losses in the West, in line with the evidence for the United States by Autor et al. (2013). A popular argument is that trade surpluses are engineered by the Chinese government through a systematic exchange rate manipulation, that is, by pegging the renminbi (RMB) to the dollar at a low value. This view has dubious foundations. Although the Chinese surplus has persisted for almost two decades, after a period of mild depreciation, the real exchange rate has been appreciating since 2003 (see Song et al. 2014). A misaligned exchange rate should feed domestic inflation, for example, by increasing the demand of nontraded goods and stimulating domestic wage pressure. However, until very recently, it does not appear as if China has experienced any major inflationary pressure—between 1997 and 2007, the inflation rate was on average about the same as in the United States. Reisen (2010) shows that the RMB was not significantly undervalued in real terms in 2008, once one filters out the Balassa-Samuelson effect. We conclude that the control of the nominal exchange rate is an unlikely first-order cause of the persistent imbalance.

Song et al. (2011) provide an alternative, structural explanation for the imbalance. The starting point is the macroeconomic identity according to which the trade surplus equals the gap between domestic savings and investments. China experienced large investment rates, but even larger saving rates. Since 1997, domestic savings have exceeded domestic investments (as a share of GDP). It is useful to decompose the savings gap between the household, corporate, and government sectors (see Figure 5). As expected, households are net suppliers of savings, whereas firms demand external resources to finance investments. The net position of the households follows a nonmonotonic pattern. In contrast, the net demand of external funds from firms (i.e., corporate investments minus savings) has been declining sharply as a share of the GDP since 1992. The government played a less important role in accounting for the aggregate savings gap.

In Song et al. (2011), a rising foreign surplus is caused by the fact that DPEs borrow less from banks. As their growth crowds out SOEs during the transition, the banks become awash in cash and must invest in foreign bonds. To see this mechanism more formally, consider the banks’ balance sheets in the simplified version of Song et al. (2011) outlined above. Let \( B \) denote banks’ purchases of foreign bonds and \( K_F \) denote the loans to F-firms. As F-firms are entirely bank financed, this is also the investment level of F-firms. Bank deposits equal the savings of the workers. Savings and investments of E-firms do not feature in the banks’ balance sheets, as managers do not hold deposits, nor do banks make loans to E-firms. Hence, equating banks’ assets and liabilities yields

\[
\frac{K_{Ft} + B_t}{\text{assets}} = \frac{\beta}{1 + \beta} w_{t-1} N_{t-1} \frac{\text{deposits}}{N_{t-1}}.
\]

(7)

As the employment share of F-firms declines during the transition, the demand for loans to finance their investments falls too. This reduces the total demand for funds from local investors, as E-firms are borrowing constrained. However, workers’ wages are unaffected by the transition, so aggregate deposits do not depend on the share of E-firms during the transition. Consequently, banks must shift their portfolios toward foreign bonds (\( B \)), causing a growing foreign surplus (Figure 3e). In the model, the accumulation of reserves can be viewed as a growing gap between

---

14This effect is partially offset by more jobs being created in export-oriented industries. Dauth et al. (2012) find that the positive employment effect dominates in Germany.
banks’ deposits and loans to domestic agents. This is precisely what we see in the data (see Figure 4), in which the growing gap between deposits and loans tracks closely the accumulation of foreign reserves.¹⁵

This mechanism is reminiscent of the empirical pattern of corporate demand for external funds (i.e., investments minus savings) in Figure 5, which is an updated version of figure 2 in Yang (2012). As the transition progresses, internally financed DPEs replace externally financed SOEs, shrinking the gap between corporate investments and savings, consistent with Figure 5.¹⁶ The theory is also consistent with the pattern of the allocation of savings and investment across regions in China. Song et al. (2011, table 1, p. 207) document that the gap between savings and investment is positively correlated with the private employment share in manufacturing, which in turn is positively correlated with productivity growth. Thus, a version of Gourinchas & Jeanne’s (2014) allocation puzzle holds across Chinese regions. The findings are confirmed by the more formal analysis by Cudre (2013), who finds that provinces with faster TFP growth have higher investment wedges (hence lower investments) and lower saving wedges (hence higher saving).¹⁷

In the theory outlined above, the foreign surplus is caused by the inability of underdeveloped financial markets to channel savings to highly profitable domestic investment opportunities (for related mechanisms, see Buera & Shin 2009, Sandri 2010, Angeletos & Panousi 2011, and

---

¹⁵Song et al. (2011) also predict a rising aggregate saving rate (see Figure 3f). The reason is that the young, who are the net savers, acquire over time a larger share of aggregate income: The old capture a share $\alpha$ of the income of F-firms and a share $\alpha(1-\phi)$ of the income of the E-firms, and the income share of E-firms grows over time.

¹⁶Empirically, a significant part of the investments of SOEs is internally financed, whereas in the model, the SOEs have no retained earnings. For the model to be consistent with the evidence, SOEs finance a larger share of their investments through external loans than do DPEs. This is true empirically (see Song et al. 2011, figure 4, p. 210).

¹⁷Readers are also referred to Cudre & Hoffman (2013) for a more detailed analysis of the determinants of net exports for Chinese provinces.
Matsuyama 2012). These contributions are related to recent theories of global imbalances emphasizing the shortage of assets in emerging economies with high saving rates (e.g., Bernanke 2005, Caballero et al. 2008, Mendoza et al. 2009, Gourinchas & Rey 2014). They emphasize that financial markets in emerging economies are underdeveloped and fail to offer savers assets that provide risk sharing and stores of value.

Global imbalances reflect financial flows. Gourinchas & Rey (2014) show that these flows are much larger than flows of foreign direct investment (FDI) to and from emerging economies. This is true also for China (Song et al. 2014). One potential reason is that there are tight controls on cross-border portfolio investments. Holmes et al. (2013) propose a complementary mechanism. They document that China is pursuing a quid pro quo policy of granting market access in exchange for gaining access to know-how and technology on which Chinese firms can build. Such an exchange could increase competition down the road for a potential investor, and this makes Western firms reluctant to do FDI in China. Holmes et al. (2013) argue that this policy can explain why the gross FDI flows to China are so small.

3.4. The Savings Puzzle

China’s massive foreign surplus points to another puzzling aspect of China’s transformation: its huge propensity to save. The aggregate saving rate has been above 35% of GDP since the 1980s and has risen further after 2000, exceeding 50% in recent years (see, e.g., Horioka & Wan 2007, Ma & Yi 2010, Yang 2012). Household saving as a fraction of disposable income has increased from 16% in 1990 to over 30% nowadays. Standard theories emphasizing consumption smoothing in contrast would predict that China should have low and decreasing saving rates. Why have Chinese households saved so much, given the fast income growth observed since the 1980s?

Chamon & Prasad (2010) and Song & Yang (2012) document an additional puzzling aspect of household saving in China. The younger and older households have higher saving rates than do the middle-aged households, so the age profile of saving rates is U-shaped.\(^\text{18}\) This is a unique feature of China, even relative to other developing countries. Moreover, it contradicts the standard life-cycle model, in which savings are large for middle-aged and small for young and old households (Modigliani & Cao 2004), given a standard increasing age-earnings profile.

The high saving rate of Chinese households and the anomalous age profile associated with it have attracted a lot of recent research. Here we review some of the studies. Chamon & Prasad (2010) document a large increase in saving rates for all cohorts and age groups. Thus, a composition effect arising from a changing age structure is unlikely to have played an important role. Instead, the rise in savings must result from factors associated with the transition to a capitalist economy. The authors argue that a key driver of high savings is the shifting burden of health and education expenditures from the state to individuals, a change that has induced young households to save for their children’s education, and elderly households to save for retirement and health care needs. The end of the so-called iron rice bowl had a particularly large effect on the savings of the generations more directly affected by the reforms (i.e., those who were in their forties and fifties in 1990). Finally, Chamon & Prasad (2010) argue that the large-scale privatization of residential houses, which triggered an increase in the ownership rate from 17% in 1990 to 86% in 2005, played an important role. The reform increased house prices and the number of potential

\(^{18}\)The findings that saving rates are falling with age for younger households are challenged by Coeurdacier et al. (2013), who argue that they may reflect an aggregation bias associated with the presence of multigenerational households. For instance, the saving rate of middle-aged household heads may be underestimated if they cohabit with younger adults or elderly who have lower savings.
buyers (see Wang 2011, Liu et al. 2013). Given the down payment constraints and financial frictions, savings increased.19

The precautionary motive argument implies that saving rates should increase with income uncertainty. Based on the methodology of Storesletten et al. (2004), Fang et al. (2010) show that Chinese households face more severe, persistent risk than do US households and hence save more. There is a concern with an explanation based solely on precautionary motives: Although household saving rose sharply after 2000 (see Figure 5), it is difficult to point to a large empirical increase in individual risk during this period. For instance, pension coverage increased over that decade.

Song & Yang (2012) argue that the increase in the saving rate, and in particular the U-shaped age profile of savings, can be accounted for by technological factors. They document that the growth rate of the entry wage across cohorts has been higher than the average wage growth in China, whereas the age profile of earnings has become flatter for later cohorts. The life-cycle model implies that young agents should save more when the age-earning profile is less steep. Using a quantitative multiperiod overlapping-generations model, they show that this mechanism can account for an increase of approximately 10 percentage points in the average saving rate. Their mechanism can also explain the increase in the saving rate observed for the 25–45 age group over the period 1992–2007.

An important policy change that may have affected saving behavior is the sequence of family planning policies introduced in the 1970s, culminating with the 1979 one-child policy. This policy imposes draconian sanctions on urban couples who have more than one child and rural couples who have more than two children, with exemptions for special groups and ethnic minorities. As a consequence, the total fertility rate (TFR) fell sharply. For instance, in urban areas, the number of surviving children fell from 3.2 in 1970 to less than 1.3 in 1982. Banerjee et al. (2013), Choukhmane et al. (2013), and Zhou (2013) argue that the policy’s introduction may have increased savings because it reduced the number of children who can potentially provide old-age transfers when parents retire.

Banerjee et al. (2013) focus on the savings of retirees over 65, using data from the China Health and Retirement Longitudinal Survey. They estimate the effect of family planning policies by exploiting the earlier policies introduced in 1972 under Chairman Mao, which encouraged an increase in birth spacing of three to four years. Even though these policies did not literally restrict the number of children ever born to a woman, they had a large negative effect on the number of children per woman because the subsequent introduction of the 1979 one-child policy came as a surprise. The parents who in 1979 had planned to have many children but had delayed the timing of the births in compliance with the spacing policy were unexpectedly barred from having more children. The authors find that parents who had their first child after 1972 (the treatment group) have on average significantly fewer children than those who had their first child before 1972 (the control group). The reduction in the number of children is smaller for parents in the treatment group whose first child was a son—likely because many parents who already had a son decided voluntarily not to have a second child, so the policy had a weaker effect on this subgroup. As far as savings are concerned, Banerjee et al. (2013) find that on average, families in the treatment group have a saving rate 5.7 percentage points higher than parents in the control group. The difference is almost entirely accounted for by families whose first child is a daughter born in 1972 or later. For these families, saving rates increased by 9.7 percentage points, whereas there is no significant

19Using the Urban Household Survey, Brugiavini et al. (2013) document that the saving rate of households who bought a house after 1998 at market price was significantly higher than that of the rest of the population (see also Wang & Wen 2010).
effect of the 1972 family planning policies on families whose first child is a son. The authors interpret this finding as related to the Chinese tradition that sons provide more support to elderly parents than do daughters. Thus, the policy had its strongest effect on the savings of families who had a daughter and were unexpectedly barred from having another child. These couples could not expect much help from their child and therefore save more during retirement.

Choukhmane et al. (2013) study the effects of the one-child policy exploiting data from the Urban Household Survey. They compare the saving rates of households who had twins after the introduction of the one-child policy with those of households without twins. Having twins is not subject to legal sanctions and is a credible source of exogenous variation. Households with twins had between 3.5 and 8.5 percentage points lower saving rates than households without twins, depending on the sample and the definition of consumption. In their data sample, children are still residing with their parents. Thus, as the authors acknowledge, part of the effect arises from a pure expenditure channel: Consumption increases when there are more mouths to feed.

Zhou (2013) focuses on the savings of the children. She documents that single children have significantly higher saving rates than do children with brothers. She argues that this is a result of siblings, and especially brothers, providing implicit risk sharing and potential for sharing the future burden of elderly care for their parents. Thus, by reducing the number of siblings, the one-child policy increased the savings of young households.

Looking at a different channel, Wei & Zhang (2011) argue that the increase in savings is related to the growing sex imbalance at birth. The sex ratio (i.e., the number of men per woman) has risen dramatically over the past three decades in China, from 106 boys for 100 girls in 1980 to 120 boys per 100 girls in 2005. This results from the joint effects of the one-child policy, the traditional preference for a male offspring, and the access to selective abortion technology. These changes stiffened the competition among boys in the marriage market. Wei & Zhang (2011) argue that, in response, households with a son will increase their savings to make their boys more competitive in the marriage market. In support of this theory, they document that households with a son have a higher average propensity to save than households with a daughter. Moreover, the savings of households with a son increased more in regions with a greater sex imbalance. Finally, saving rates tend to be higher in regions and years with a greater local sex imbalance.

The studies reviewed above suggest that demographic changes can be important for savings. However, changes in fertility may have different effects on savings at different stages of the life cycle (see, e.g., Ge et al. 2012). Owing to data limitations, we do not yet know how the one-child policy will eventually affect savings for all age groups. In addition, the regression studies can uncover only partial equilibrium effects. Banerjee et al. (2013) show that a general equilibrium effect through the interest rate can mitigate significantly the partial equilibrium effect of fertility restrictions on saving rates.

4. THE FUTURE OF CHINA

What does the future bode? A common view in the West is that China’s growth trajectory is unsustainable, owing to the persistence of a nondemocratic institutional framework, different from those that promoted prosperity in the West (see Acemoglu & Robinson 2012). The extractive Chinese institutions can possibly sustain catch-up at an early stage of development but will eventually become a burden. According to this view, China will not escape an institution-driven middle-income trap. Others argue that China’s hybrid form of state capitalism can become a successful new model of economic growth, possibly exportable to other developing nations (Musacchio & Lazzarini 2012). Owing to the lack of historical precedents, this debate is necessarily speculative. In light of the recent economic literature, this section reviews factors
influencing China’s future economic development. We start with some impetuses and then discuss some obstacles.

4.1. Impetus for Sustained Growth

First, the mere adoption of technologies already in use abroad is likely to be a powerful engine of growth for at least another decade. China is still a relatively poor country, with an average productivity of firms and workers far below that of industrialized nations. Therefore, the potential for technological catch-up is still enormous. To draw a comparison, in 2010 the GDP per capita of Russia was more than twice that of China (Penn World Table 7.1). Given the current rates of TFP growth and fast technology adoption, it seems unlikely that China’s institutions could prevent it from reaching Russia’s current level of development.

Second, China is being transformed into an innovation-oriented economy. FDI has been a major driver of technology transfer (see Acemoglu et al. 2012). In addition, China has also invested large resources to develop a strong local innovation capability, which Holmes et al. (2013) view as complementary to the foreign technologies transfer via FDI. According to the OECD Science, Technology, and R&D Statistics 2012, the R&D investment share of GDP has increased steeply, from 0.7% in the 1990s to 1.84% in 2011. Today’s figure is close to the average for the European Union (1.94%), albeit still lower than that of the United States (2.77%). China has already surpassed rich economies such as Canada, Italy, Spain, and the United Kingdom and spends substantially more than do other emerging economies. The comparison is even more impressive if one restricts attention to industry-related R&D investments: China invests 1.36% compared with 1.66% of the United States and 1.02% of the European Union. China employs 820,000 researchers working in business enterprises, whereas the United States employs 1.1 million of them. Finally, looking at the number of patents in 2011, according to the World Intellectual Property Organization, the Chinese patent office granted 172,000 patents, compared with 225,000 and 152,000 patents granted by the US Patent and Trademark Office and the European Patent Office, respectively.20

Interestingly, although China’s growth has been so far largely export led, the recent boom in innovation appears to be increasingly driven by the expansion of the domestic market. Beerli et al. (2013) document that a massive change in domestic demand is taking place as Chinese households are lifted out of poverty and a rapidly growing middle class can afford to spend on durable consumer goods. The expected growth of the future domestic demand (which varies across different durable goods) is shown to be a significant determinant of innovation at the firm level after controlling for potential endogeneity issues. Their results are consistent with the predictions of models of directed technical change (see, e.g., Acemoglu & Zilibotti 2001, Gancia & Zilibotti 2009), augmented with nonhomothetic preferences (Boppart & Weiss 2013). According to these models, the growth in market size related to changes in income levels and distribution should predict the direction of R&D investments.

In summary, China is a very innovative economy, far ahead of any other country at a comparable development stage. Economic theories arguing the case for middle-income traps (see, e.g., Acemoglu et al. 2006) emphasize the inability of countries with rigid institutions to promote a transition from an investment-driven to an innovation-driven growth. There is no evidence indicating that China is heading toward a low-innovation equilibrium.

---

20In 2011, the number of applications filed in China was actually higher than the number of applications filed in both the EU and the US patent offices. This suggests that inventions patented in China might be of a lower quality (see Bussy 2013 for more discussion).
Third, China is making large human capital investments. The average years of education in the population over 25 is now 7.5, twice as high as in 1980 (Barro & Lee 2012). Assuming a rate of return to education of, for example, 10% per year, this increase in educational attainment implies an increase in average labor productivity of 43% owing to human capital accumulation alone. The increase in the proportion of the population with some tertiary education is even faster: 6% today against 1% in 1980. The current enrolment rate in tertiary education (27% in 2013, according to the World Bank’s World Development Indicators) implies steep increases in future educational attainment. The boom in higher education is a recent phenomenon: Between 1979 and 1995, tertiary school enrollment rates were below 2.5%, and the increasing shortage of high-skill workers was reflected in a rapid rise in the return to education. This trend was followed by a rapid expansion in higher education during the past decade, bringing the number of fresh college graduates from less than 1 million in 2001 to over 6 million in 2010 (Heckman & Yi 2012). This was accompanied by a booming number of students studying overseas: Chinese students represent today 18.2% of all international students enrolled in Organization for Economic Co-operation and Development (OECD) countries (OECD 2011). Changes in the quantity of education likely underestimate the actual human capital accumulation. During prereform times, schools at all levels emphasized ideological learning. In 2009, Shanghai came out as the leader in the PISA study, outperforming Western countries by a wide margin. Interestingly, Shanghai students’ scores were far less correlated with their socioeconomic background than in OECD countries. Thus, schools appear to be a vehicle of social mobility in China.

Human capital accumulation, investments in technology adoption, and industrial policies (e.g., SEZs) have come hand in hand with an increasing technology intensity of industrial production. Wang & Wei (2010) document that over the past decade, China has become increasingly less specialized in labor-intensive and low-value added industries, shifting its production and export structure toward high-technological sectors.

Finally, an important, yet largely unexploited potential source of future growth is the reduction of the pervasive financial frictions. In Section 3 we argue that these are responsible for severe misallocations. We believe that reforms aimed at reducing the market power of the large state banks, for instance, by allowing banks to compete in offering deposit and lending rates (so far heavily regulated), and at improving the legal system (e.g., contract enforcement, investor protection) can have large effects on productivity. The opening of the capital account and the convertibility of the RMB, currently under discussion, are an opportunity for such reforms.

An important first step in this strategy is the Shanghai Free Trade Zone (SFTZ), launched in September 2013 under the impetus of Li Keqiang’s government. The goal of the SFTZ is to become a test ground for financial reforms and to accelerate the capital convertibility and full liberalization of the financial service industry. Although many important details about the regulation of the SFTZ are yet unclear, there are reasons to believe that this initiative may become a stepping stone in the process of financial reform. The special policies applying to the zone are expected to be full interest rate liberalization, convertibility of the RMB (even for non-trade-related transactions), free entry of foreign banks, and permission for domestic banks to provide off-shore financial services. In addition, there will be neither control nor tariffs on goods entering the SFTZ from abroad, and free circulation of goods within the zone, although strict controls will apply to the flow of goods between the SFTZ to the rest of mainland China. However, there are reasons to believe that the rigid insulation of the zone may be relaxed in future. That other cities in China

---

Notes:

21The growth rate of wages of college graduates was 240% between 1992 and 2007 compared with 135% for workers with a middle school education and below (Ge & Yang 2013).
such as Tianjin and Chongqing have offered to host similar free trade zones suggests the possibility that financial liberalization will undergo a gradual expansion following the model of the industrial policy (SEZs) of the 1980s discussed above.

4.2. Obstacles and Challenges to Sustained Growth

Despite China’s potential promise, there are looming issues that could potentially stall the reform process and threaten the future economic development. We discuss here four salient factors: the rising economic inequality, the aging population, the environmental toll of rapid economic growth and its negative impact on quality of life, and the fragility of the political equilibrium.

4.2.1. Inequality and aging. Economic inequality and the aging of the population are perhaps the most acute challenges to social cohesion and the status quo in China. Fast growth has been accompanied by a rapid increase of income inequality. Starting poor but equal, China has evolved into a highly unequal society. According to the CIA World Factbook, the Gini coefficient of income has risen from 0.36 in 1992 to 0.47 in 2012, roughly comparable to that of the United States, and much higher than that of any Western European country. According to the World Top Income Database, the top 10% income share rose from 19% to 28% between 1990 and 2003. The sources of increasing inequality are manifold: age (i.e., younger cohorts are much richer than older ones), education, dispersion across regions and between rural and urban areas, resident and nonresident workers, capitalists and workers, etc. To the extent that a continued reform process—the status quo—requires social cohesion, the growing disparities pose a real risk. In absolute terms, growth has benefitted both the rich and the poor. As discussed above, the poverty rates have fallen dramatically after 1980. Over the same period, life expectancy has risen by approximately 10 percentage points, reaching 74 years in 2011. Although growth has benefitted the Chinese population overall, inequality may pose a threat to social cohesion and has become increasingly salient in the discourse of the political leadership. In a press conference held on March 14, 2012, the former Premier of the State Council Wen Jiabao declared, “I know that social inequities...have caused the dissatisfaction of the masses. We must push forward the work on promoting social equity.... The first issue is the overall development of the reform of the income distribution system.”

The looming aging of the population is another such critical issue. The total dependency ratio has fallen from 75% in 1975 to just 37% in 2010. This is a result of the combination of high fertility in the 1960s, when China’s TFR was between five and six, and the family planning policies introduced in the 1970s, culminating with the one-child policy. As a result, a very large share of the Chinese population is of working age today. The expanding share of working people has contributed to economic growth in the two past decades. However, China has now reached a turning point: The old-age dependency ratio will increase from the current 12% to 39% in 2040. The trend is likely to continue well beyond 2040, as the current TFR (estimated to be approximately 1.6; see Yi 2007) is below the replacement level. Even if the one-child policy were relaxed, it is unclear to what extent this would increase the TFR, at least in urban areas. For instance, other emerging economies that are today richer than China, such as Brazil and Russia, have low TFRs (equal to 1.8 and 1.7, respectively). Regions culturally similar to mainland China, such as Hong Kong, Taiwan, Macau, and Singapore, have TFRs of 1.2 or lower.

Some commentators warn that an aging population might harm economic growth, pointing to the experience of Japan: Future saving rates may decline (although this is unlikely to be a major issue for China); the society may turn less forward-looking and innovative; and, most importantly, the government’s future tax base, and hence its ability to finance social policies, will be
compromised. Thus, an aging population will make it increasingly difficult to mitigate China’s income disparities.

In no setting is the dual challenge of inequality and aging more evident than in China’s pension system. Historically, pension systems have been a powerful vehicle of intergenerational redistribution in Western economies, and this social contract has been a force for social cohesion. Its introduction bailed out the unlucky generations that were hit by the Great Depression and World War II. Arguably, a similar case can be made for the current elderly Chinese workers who were impoverished by the tragic experiences of the Great Famine and the Cultural Revolution. Indeed, intergenerational inequality is an important part of total inequality in China. Owing to high growth, the present value of earnings for a worker entering the labor force in 2000 is on average about six times as large as that of a worker who entered in 1970 (Song et al. 2012). Poverty among the elderly is a major social issue, especially in rural areas (see, e.g., Yang & Chen 2010, Almås & Johnsen 2012). One might object that the elderly can be bailed out by their own children. However, the traditional family insurance system relying on transfers and support from children (especially sons) is under strain, owing to the smaller number of children, increasing geographic mobility, and the decline of traditional values (see, e.g., Cai et al. 2006, Park et al. 2012). In this context, pension transfers seem critical for mitigating inequality and poverty.

4.2.2. Sharing the benefits of growth: the role of pension reform. Given these issues, it is natural that the pension system is a key policy issue for China. We now review China’s current pension system and then analyze how it should be changed, in light of our preceding discussion on inequality and aging.

China has an urban pension system, originally introduced in 1986 and then reformed in 1997. Rural residents earn no pensions, although a limited rural pension system has been introduced recently. Prior to 1986, urban firms were committed to paying pensions to their retired employees. This system ceased to be viable in a market economy with firm and worker turnover. The 1986 reform transferred responsibility for pensions to local municipalities. However, private firms were typically evading contributions, and many urban workers did not accumulate pension rights. Municipalities came under financial distress and had to be bailed out by the central government. The 1997 reform reduced the generosity of pensions and strengthened the enforcement of contributions. Subsequently, the coverage of the system has now risen from 44% in 1992 to over 60% today. However, the current system is not financially sustainable. Song et al. (2012) find that given the demographic outlook of China, the present value of the future contributions falls short of the present value of the promised pension payments. They estimate that if one were to achieve sustainability by adjusting the replacement rate as of 2013, a permanent reduction to 40% (down from the current 60%) would be required.

There are several alternative ways to make the pension system sustainable, each involving a different extent of intergenerational redistribution. Song et al. (2012) study the welfare implications of a range of sustainable pension reforms, assuming that the economy first goes through a period of fast wage growth (transition) and then experiences a slowdown (steady state). To evaluate welfare consequences of intergenerational redistribution, they introduce a planner who cares about all present and future generations, discounting the future generations’ utilities geometrically. The planner has two conflicting objectives: to minimize tax distortions and to achieve the desired intergenerational redistribution. The weight on these two objectives depends on the social discount rate: the more forward-looking the planner, the lower his or her drive to redistribute consumption toward the earlier generations.

With the aid of a calibrated multiperiod version of the model of Section 3, augmented with endogenous labor supply, Song et al. (2012) analyze the welfare effects of a variety of financially
sustainable reforms, including a fully funded reform that eliminates any intergenerational transfer (and tax distortions), and a pure pay-as-you-go system. They find that even a planner with an annual social discount rate as low as 0.5% would strictly prefer the pay-as-you-go reform to a fully funded reform, or to a sustainable immediate reduction in the pension benefits. The reason is that the pay-as-you-go system delivers generous pensions (in excess of the current 60% replacement rate) over the next 30 years, owing to the demographic structure characterized by a low old-age dependency ratio. Even though the future generations receive very low pensions, these generations have very high wages and can save for their old age. Thus, the social planner is satisfied with a system of intergenerational transfers that shifts resources away from them to bail out the initial generations of poor workers. Interestingly, the result hinges on the forecast that Chinese wages will continue to grow for some time (according to the forecasts of Song et al. 2012, wage growth will exceed an annual 5% until 2030, excluding the effect of human capital accumulation). If wage growth were constant at a 2% annual rate, the planner would instead prefer to switch to a fully funded system.

These normative predictions run against the popular argument that reforming the pension system in a prefunded direction is the appropriate response for emerging economies with an aging population (see, e.g., Feldstein 1999, Feldstein & Liebman 2006, Dunaway & Arora 2007). They are instead broadly in line with the policy recommendations of Barr & Diamond (2008). Song et al. (2012) provide a rationale for using a temporarily unbalanced pension system to bail out the poor generations that are currently middle-aged and older. This can be a vehicle of social cohesion in China, especially if it is extended to cover all Chinese citizens. Carrying out an immediate fiscal consolidation of the pension system would require substantial government savings today to finance future transfers. Given the large wealth that China has already amassed (e.g., in the form of foreign reserves), this does not seem to be a top priority. Alternatively, this wealth could be used to finance policies and institutions aimed at reducing inequality and making the growth process more inclusive. As China develops, the citizens’ demand for more extensive welfare policies is likely to grow. A large share of citizens, such as the rural and the nonresident migrant population, can be expected to want access to social insurance from which they have been largely excluded in the past. There is, in our view, an urgent need for more quantitative work evaluating the costs and benefits of social reforms to attain an efficient institutional design.

### 4.2.3. The environmental disaster: pollution for promotion

One of the big open questions for China is the bias of its growth process toward quantity relative to quality. Environmental degradation casts a shadow over China. Kahn & Yardley (2007) write, “Chinese cities often seem wrapped in a toxic gray shroud. Only 1% of the country’s 560 million city dwellers breathe air considered safe by the European Union.” Medium-size cities such as Linfen and Tianying lead the sad ranking of the most polluted cities worldwide, owing to coal and heavy metal contamination, respectively. Water pollution is an equally severe problem (Shapiro 2012). China relies on dirty water and polluted air, which have a significant impact on public health and the environment.
technologies to a larger extent than do other countries of a similar development level (Vennemo et al. 2009). Why has the problem grown so rampant?

According to Jia (2012), the answer is the system of promotion incentives within the CCP that induces local political leaders to disregard environmental considerations to achieve maximum growth. The promotion of provincial governors wishing to climb the Party’s hierarchy ladder hinges on the growth performance of their province. This inhibits politicians from taking or enforcing measures that would limit the use of cheap polluting technologies.

To investigate the question more formally, Jia (2012) builds a model whose main prediction is that if a provincial governor stands a serious chance of promotion, he or she will disregard environmental concerns. The theory implies that if a local leader acquires a previously nonexisting political connection with a senior politician that enhances career opportunities, the leader’s region will experience an increase in both the use of dirty technologies and economic growth. She tests this theory by focusing on the connections between provincial governors and members of the Politburo Standing Committee, who are responsible for the promotion of provincial governors to higher positions within the CCP. To avoid endogeneity issues, she assumes that a governor is connected to a Standing Committee member whenever they were work colleagues, studied at the same university, or originated from the same province. Such contacts occur early and are not affected by subsequent steps in the political career. She estimates how the performance at the provincial level changes when a politician to which the provincial leader is connected ascends to the Politburo Standing Committee. She finds that when a governor becomes connected, industrial growth increases in his or her province relative to the rest of China. However, the environmental toll is heavy: Water contamination from both industrial waste and air pollutant emissions increases by approximately 25% and 10%, respectively.

Jia’s (2012) findings support the hypothesis that career concerns can help explain why China is so heavily polluted today. This suggests that political reforms should focus not only on passing new laws and financing programs to stop the environmental degradation, but also on changing the informal rules and incentives within the CCP—establishing good environmental standards as an explicit measure of political success. Another implication is that a more environmentally balanced approach may require some sacrifice in terms of growth rates. If such policy changes can avert environmental disasters, they may improve welfare for millions of people.

4.2.4. State capitalism and political reforms. Political incentives also are at the core of another topical question: Will economic growth bring about democratization, or will the current political system be resilient to changes? If so, will political distortions curtail economic development?

In a recent paper, Y. Wang (2013) addresses these important questions with the aid of a theory in which a self-interested political elite controls state-owned firms and can impose taxes on private firms. To stay in power, the elite must shore up sufficient political support to avert a democratic revolution (as in Acemoglu & Robinson 2005). The political elite achieves this goal through a divide-and-rule strategy that involves distorting the economic equilibrium so as to grant sufficient privileges to its supporters.

Y. Wang (2013) derives conditions under which the economic transition will eventually trigger democratization. When these conditions fail, the elite retains power perpetually and continues to impose distortions to the economic allocation that may eventually harm economic development.

25There is a growing empirical literature studying the career incentives of Chinese provincial leaders. Persson & Zhuravskaya (2013) study the effects of the social ties between provincial leaders and provincial elites on governance. Jia et al. (2013) study the complementarity between connections and performance.
Y. Wang (2013) formalizes the argument with the aid of a two-sector growth model along the lines of Song et al. (2011) (reviewed in Section 3). The novel feature is that the political elite owns the capital stock of F-firms and extracts a surplus by taxing the more productive E-firms owned by middle-class entrepreneurs. To avert a democratic revolution, the elite must secure the support of a sufficiently large share of workers. To this aim, the elite creates a dual labor market, in which the F-firm workers are paid a wage premium and become the base of support of the political system. In contrast, the wage paid by E-firms is competitive. Under democracy, the median voter (i.e., the workers) taxes all firms and redistributes the revenue as lump-sum transfers. To avert a revolution, the elite must satisfy two constraints: A minimum share $\pi$ of the workers must be employed by F-firms, and these workers must earn a wage in excess of what they would earn under a transition to full democracy: $w^F_t \geq w_{t}^{dem}$, where $w_{t}^{dem}$ is the equilibrium wage (including transfers) under democracy. There is no guarantee that a nondemocratic equilibrium exists. For instance, if the productivity or the share of total capital of F-firms is too low relative to that of E-firms, the elite becomes unable to satisfy both constraints and must concede democratization.

Figure 6 illustrates the case of a nondemocratic equilibrium for a given private and public capital stock. The two schedules represent the marginal product of labor in F-firms (downward-sloping curve) and in E-firms (upward-sloping curve). The undistorted competitive equilibrium yields employment $L^D_F$ in F-firms and a common wage rate $w^D$. This allocation cannot be an equilibrium under nondemocracy, as workers would then support a revolution. The nondemocratic equilibrium yields instead $w^F = w^{dem}$ and $L^{ND}_F \geq \pi L$.

The theory is consistent with the puzzling observation that the middle class, often regarded as the driving force of democratization in the development process, appears to support the non-democratic system. In Y. Wang’s (2013) model, the nondemocratic equilibrium appeals to the workers in the state sector, and also to the entrepreneurs who, owing to the labor market distortion, can hire workers at a lower wage. The sole losers are the E-firms workers who earn lower wages ($w^E_t$) than under democracy and receive no transfers. The prediction of the theory conforms with the evidence from Chen & Lu (2011), who document that state sector employees on average are less supportive of democratic values. Interestingly, employment status (state or private sector)
is a stronger predictor of the attitude toward democracy than is party membership. Also consistent with the theory, Ge & Yang (2013) document the existence of a significant wage premium for observationally equivalent workers employed in the state sector.

Capital accumulation in the private sector has two opposite effects for the elite. On the one hand, it increases the revenue of E-firms and the tax base from which the elite can extract a surplus. On the other hand, as the hiring potential of private firms increases, it threatens the political employment constraint $I_{ND}^E \geq \pi L$. Thus, although partial privatization initially is welcome by the elite, when the E-sector becomes sufficiently large, the elite has an incentive to curtail private employment and to increase F-firm investments. One such policy is to impose tight credit constraints on E-firms, so as to slow down their growth and give a significant advantage to F-firms in the credit market. Thus, the theory provides a rationale for strategic distortions of financial markets, such as those documented in Section 3. It also provides a rationale for increasing investments (again, for strategic reasons) in the state sector as the private sector grows, consistent with the evidence of Hsieh & Song (2013).

At an early stage of the transition process, state capitalism speeds up growth by suppressing wages and favoring private sector growth. However, at later stages, survival of the regime requires distortions to sustain sufficient SOE employment. In some cases this can lock the economy in a middle-income trap. In other cases, the equilibrium features a serendipitous transition to democracy.

5. CONCLUSION

In this article, we review some central issues in the recent economic development of China. We emphasize, with the aid of a model, the transitional nature of China’s growth process over the past three decades. China now faces a dilemma: The scope for growth driven by reallocation, along the lines of the model in Section 3, is diminishing, making future growth more dependent on local innovation and human capital. Owing to its large investments in R&D and education, China is likely to get a soft landing. Although growth may slow down, we see no indication that China will get stuck in a middle-income trap. Still, the current model of state capitalism relies on important distortions. It is an open question whether the political elite has the incentives to overcome such inefficiencies and complete the reform process (e.g., a further reduction of the state’s role in economic activity), as this may trigger an increase in the demand for political changes. In the future, fostering social cohesion and averting environmental disasters will be critical policy issues.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Simon Alder, Adrien Bussy, Sebastian Ottinger, Maria Saez Marti, Lin Shao, Zheng Song, Roman Studer, and Yikai Wang for useful comments and research assistance. Financial support from the European Research Council (ERC Advanced Grants IPCDP-229883 and MACROINEQUALITY-324085) is gratefully acknowledged.

26In terms of Figure 6, suppose that the E-firm (upward-sloping) schedule is shifted to the left. Then it may become impossible for E-firms to satisfy simultaneously the two constraints $u^F \geq u^{dem}$ and $I_{ND}^E \geq \pi L$. Thus, if the share of the capital stock of E-firms is too large, the elite cannot retain power.
LITERATURE CITED


Boppart T, Weiss F. 2013. *Non-homothetic preferences and industry directed technical change*. Unpublished manuscript, Univ. Zurich


Feldstein M, Lieberman J. 2006. Realizing the potential of China’s social security pension system. *China Econ. Times,* Febr. 24


Hsieh CT, Song Z. 2013. *Grasp the large, let go of the small: the transformation of the state sector in China.* Unpublished manuscript, Booth Sch. Bus., Univ. Chicago


Matsuyama K. 2012. *Institution-induced productivity differences and patterns of international capital flows*. Unpublished manuscript, Northwestern Univ., Evanston, IL


