Are China and India Backward? Evidence from the 19th Century U.S. Census of Manufactures

Nicolas L. Ziebarth*

January 21, 2011

Abstract

Hsieh and Klenow (2009) argue that a large fraction of aggregate TFP differences between the U.S. and the developing countries of China and India can be explained by capital misallocation. Their interpretation is that this misallocation is due to institutions and policies that distort resources away from productive firms in these developing countries. Using the U.S. Census of Manufactures from the late 19th century, I find that the level of dispersion in these modern, less developed countries is very similar to that in the U.S. at this time. What these countries share are not similar institutions rather similar levels of economic development. The institutions of the U.S. at this time were much better than India or China in terms of protecting property rights and allocating resources. This suggests that the Hsieh-Klenow measure of imperfections is not related to institutions but simply the level of development. I apply their accounting procedure to the U.S. and find that almost 15% of manufacturing TFP growth between 1890 and 1997 can be attributed to a more efficient intra-industry allocation of resources.

1 Introduction

In an accounting sense, most of the differences between more and less developed countries are due to differences in productivity rather than in the inputs to production such as human or physical capital, labor, or natural resources (Hall and Jones, 1999; Hsieh and Klenow, 2010). So what explains differences in productivity? The leading answer is that these differences arise due to differences in economic institutions and policies across countries. But the content of this explanation depends on the extent to which we can measure which institutions matter and how they affect productivity.

Hsieh and Klenow (2009) address one such channel through which institutions can affect productivity, which is the misallocation of resources from productive firms to unproductive ones. They construct a model that allows them to measure how much total factor productivity (TFP) is reduced by misallocation and apply the method to modern data from India, China and the U.S. They

*Northwestern University. E-mail address: nlz@u.northwestern.edu. I thank Jonathan Parker for encouragement and detailed comments. Jeremiah Dittmar, Matthias Kehrig, Joel Mokyr, and Chris Vickers offered useful feedback. Alex Field provided some unpublished data regarding TFP in the U.S. manufacturing sector.

find much higher levels of misallocation in the developing countries compared to the U.S and argue that the relevant difference between the U.S. and the developing countries is better institutions for allocating capital in the U.S. These include the absence of large state-owned enterprises in China and onerous licensing requirements as in India to name a couple. In the process, they strengthen the case for the role that institutions play in efficiently allocating resources and thereby raising aggregate productivity.

In this paper, I apply the methodology of Hsieh and Klenow (2009) (HK) to another economically developing country: the late 19th century U.S. The method developed by HK involves backing out from data on value added and factor inputs a “wedge” that reflects the difference between the actual and efficient outcomes. The idea in their model is that *marginal revenue products* as compared to marginal physical products should be equalized across firms and any deviations from that they take as evidence for misallocation.

The data I use is from the Census of Manufactures (COM) collected by the federal government and digitized by Attack and Bateman (1999) (AB). The data are samples from the 1850, 1860, 1870, and 1880 enumerations and include a wealth of information on revenues, input costs, capital stocks, and laborers. I find levels of factor misallocation in the U.S. during the 19th century essentially identical to those of modern day India and China. This finding casts doubt on the poor institutions explanation of factor misallocation in India and China, since the U.S. in the 19th century had relatively strong institutions compared to present day India and China. Compared to the other countries, the U.S. was and still is an excellent economic environment with limited government interference and strong property rights protection.

Rather the evidence suggests that reallocation is a “natural” part of the development process. Poor countries start out with a poor allocation of resources along with low levels of capital. Then over time, capital is reallocated to better uses. More capital is accumulated and these processes feed on each other and fuel growth. Instead of viewing misallocation as essentially fixed by policy, what is needed is a model that explains the changes in dispersion over the course of development. This would be similar in spirit to how the neoclassical growth model shows how the capital stock evolves and grows in the course of an economy’s development.

While questioning the causal interpretation of HK’s results, the framework still provides a useful accounting tool much like the growth accounting methodology of Solow. I use the HK

methodology then to calculate that the U.S. in 1850 and 1860 could have more than doubled measured manufacturing TFP by moving to the allocation of resources enjoyed by the U.S. in 1997. From 1870 to 1997, the gain while lower was still approximately 65%. Putting this in the context of the long run development of the American economy, 15% of the growth in manufacturing TFP from 1870 to 1997 can be attributed to a superior allocation of resources *within* industries. This calculation illustrates the major role reallocation plays in economic history and development more broadly.

This paper builds on a number of different strands in the literature. First, it relates to other papers using this COM dataset to understand the development of the American economy. Attack et al. (2004b) relate growing inequality in wages between 1850 and 1880 to the growing concentration of large firms, and Attack et al. (2004a) study the increasing capital-labor ratio in a variety of industries over this time period. Hutchinson and Margo (2006) ask how the Civil War affected Southern manufacturing. They show unsurprisingly that capital-labor ratios plummeted in the South after the war as a major chunk of their capital stock was destroyed. Finally, Attack et al. (2008) link the manufacturing data to information on railroad networks. They find evidence that railroads had a large positive impact on the size of establishments.

Second, this paper relates to the influential view of North (1989) and Acemoglu et al. (2005) that institutions are the fundamental cause of economic growth. For empirical support, Acemoglu et al. (2001) relate institutional differences to quasi-random differences in settler mortality. Using mortality as an instrument, they show that institutions have a very strong positive effect on long run growth. More recent work by Acemoglu and Robinson (2009) has outlined a theory of how institutional choice interacts with the power of elites. In the process, the authors offer a complete theory of why some countries reform and grow, while others stagnate.

However, this work on institutions has not focused on the precise mechanism through which institutions impact growth. So, thirdly, my paper relates to the literature that has arisen which sees the link between institutions and growth in TFP. Parente and Prescott (2000) focus on institutions that allocate monopoly rights and increase the cost of adopting technology. These are what they call “barriers to riches” that they argue explain the differences across countries. Restuccia and Rogerson (2008) focus on how “distortions” at the microlevel can generate large losses in aggregate

TFP due to misallocation. What exactly those distortions or wedges are is meant to be vague.¹ A particular policy they consider (also studied by Guner et al. (2008)) is a restriction on the size of a firm. Because in an efficient allocation, factor products are equalized, it must be the case that large firms have high levels of TFP. Hence, a size restriction falls mainly on more efficient firms, and therefore has a very large negative effect on overall TFP.² Lagos (2006) develops a theory of TFP aggregating up from the microlevel with a focus on search frictions originating in the labor market. In this setting, policy interacts with the choice to keep open certain production units in the face of idiosyncratic shocks to productivity. Policies that attempt to smooth out the response to these shocks through, say, firing costs can hurt TFP by forcing firms to continue operating unproductive technologies.

While the papers just mentioned have been mainly theoretical workers, other papers have used microdata to directly assess the degree of capital misallocation. Papers such as Banerjee and Munshi (2004); Midrigan and Yu (2009); Petrin et al. (2009) have documented large amounts of dispersion in productivity across firms in a variety of countries, which provides *prima facie* evidence for capital misallocation.³

2 Methodology

This section describes the methodology in Hsieh and Klenow (2009) for recovering the distribution of distortions. We consider two distortions: (1) a “capital” wedge τ_{Ksi} and (2) an “output” wedge τ_{Ysi} where s is the industry firm i is in. One can think of these as taxes in an abstract sense that distort outcomes from the efficient allocation. This wedge methodology has also been used in business cycle accounting of Chari et al. (2007) (See as well Hall (1997)).

Output in industry s is a CES aggregate of M_s differentiated products

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad (1)$$

¹HK make a similar assumption.

²These sorts of policies have been in place in India and are gradually being relaxed.

³To be sure, the idea that misallocation plays a major role in generating aggregate TFP dispersion is not universally shared. For example, Midrigan and Yu (2009), using Korean and Columbian manufacturing data, find that once firm growth rates are controlled for, the costs of dispersion on aggregate TFP appear much smaller.

where σ is the elasticity of substitution. Each firm is a monopolist in its differentiated product. Profits are given by

$$\pi_{si} = (1 - \tau_{Y_{si}})P_{si}Y_{si} - wL_{si} - (1 + \tau_{K_{si}})RK_{si} \quad (2)$$

Here $P_{si}Y_{si}$ is revenue net of input costs i.e. value added. Firm output is produced according to standard Cobb-Douglas production function⁴

$$Y_{si} = A_{si}K_{si}^{\alpha_s}L_{si}^{1-\alpha_s} \quad (3)$$

where α_s is an industry specific capital share. The variables stressed by HK following Foster et al. (2008) are

$$TFPQ_{si} = A_{si} = \frac{Y_{si}}{K_{si}^{\alpha_s}(wL_{si})^{1-\alpha_s}} \quad (4)$$

$$TFPR_{si} = P_{si}A_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_s}(wL_{si})^{1-\alpha_s}} \quad (5)$$

Equation 4 is the standard Solow residual, which measures how productive a firm is in terms of physical output. Equation 5 measures how productive a firm is in terms of revenue or monetary output. Their analysis shows how measures of TFPR relate to the wedges in output and capital. Precisely, HK show that

$$TFPR_{si} \propto \frac{(1 + \tau_{K_{si}})^{\alpha_s}}{1 - \tau_{Y_{si}}} \quad (6)$$

In turn, industry TFP can be expressed as

$$TFP_s = \left(\sum_{i=1}^{M_s} \left[A_{si} \frac{\overline{TFPR}_s}{TFPR_{si}} \right]^{\sigma-1} \right)^{\frac{1}{\sigma-1}} \quad (7)$$

where $\overline{TFPR}_s$ is the geometric average of the average marginal revenue product of capital and labor in sector s . The average is weighted by the share of sectoral value added by firm i . What is clear from this expression is that the relevant summary statistic for efficiency losses due to misallocation is the some measure of the dispersion of TFPR. In the case where $A_i, \tau_{K_{si}}$ are lognormally distributed,

⁴Note that there are no dynamics in this setup as compared to Midrigan and Yu (2009). Along with that comes an implicit assumption that the distribution for A_{si} is taken to be exogenous. Hence, this model leaves no room for reallocation and productivity improvement through entry and exit.

HK show that industry TFP is given by

$$\log TFP_s = \frac{1}{\sigma - 1} \log \left(\sum_{i=1}^{M_s} A_{si}^{\sigma-1} \right) - \frac{\sigma}{2} \text{Var}(\log TFPR_{si})$$

This expression illustrates the importance of both the elasticity of substitution σ and the variance of TFPR. Holding the variance fixed, higher values of σ meaning less substitution between goods decreases industry TFP. Similarly, fixing σ , more dispersion also depresses industry TFP. In the general case, any deviation of a particular firm's TFPR the industry average decreases industry-level TFP.

I follow HK who set the remaining key parameters of σ, R based on other work. In particular, HK chose $\sigma = 3$ and $R = .1$. Of course, these key parameters might have been different in the 19th century. There is some evidence that the cost of capital was higher at this time (Bodenhorn and Rockoff, 1992). For σ , HK note that the international trade and industrial organization literature place its value somewhere between 3 and 10 using modern data. It is difficult to say to what extent these values are applicable to the 19th century. HK and I make the conservative choice of $\sigma = 3$.

That being said, different parameter choices have limited implications on the inference conducted here. First, choosing a different value for R only affects the average capital distortion but not the relative comparison. HK make this point clear in their paper as well. In addition, the value of R does not affect the efficiency calculation I undertake. The choice of σ is important only in terms of the TFP accounting exercise across countries. The closer to 1 the parameter σ is, the smaller loss in overall TFP for a given level of misallocation. However, this parameter does not affect the basic measures of dispersion I calculate only how they translate into overall TFP.

Following HK, I use the capital shares from the NBER Productivity Database inflated by a factor of 3/2. I aggregate those capital shares up to the relevant 3-digit SIC code and merge that with my sample.⁵ Of course, it may be that the capital share for a given industry changes over time so the share in the late 20th century has little information. I have also repeated the calculations simply setting the capital share $\alpha_s = 1/3$ for all s with little effect on the results. I then calculate

$$1 + \tau_{Ksi} = \frac{\alpha_s}{1 - \alpha_s} \frac{wL_{si}}{RK_{si}} \quad (8)$$

⁵AB mapped the products produced by the firms in their sample into SIC codes

$$1 - \tau_{Y_{si}} = \frac{\sigma}{\sigma - 1} \frac{wL_{si}}{(1 - \alpha_s)P_{si}Y_{si}} \quad (9)$$

$$A_{si} = \kappa_s \frac{(P_{si}Y_{si})^{\frac{\sigma}{\sigma-1}}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}} \quad (10)$$

I trim the 1% tails of TFPQ and TFPR like HK and then recalculate industry averages. The variables reported are log differences between a particular firm value and the industry average by year.

3 Comparing the Economies

Before turning to the empirical results, I first compare the economies along a variety of development and policy dimensions. Along almost any policy dimension, the U.S. in the 19th century had a far better economic framework for allocating resources in relation to modern China and India.

3.1 Basic Statistics

In Table 1, I report some aggregate statistics for the 3 countries over the time periods in question. While clearly there are some differences between these countries in terms of level of income and growth rate, they are much more similar to each other than to the present day U.S., which HK use as the “control” country in their study. The countries also appear similar on other measures of development, including fraction of the labor force in agriculture and the urbanization rate. A fair characterization of the U.S. in the late 19th century is that of a developing country. In fact, a cursory view of the data would suggest that misallocation should be lower irrespective of institutional differences. Note that the U.S. at this time is still richer than China or India today. If there is a negative reduced form link between misallocation and development, then the fact that the U.S. is richer at this time would suggest a lower level of misallocation, all else equal.⁶

One strong similarity is the pace and scope of industrialization and its handmaid, urbanization. Atack and Passell (1992) argue that industrialization was already spreading in the antebellum years with the rise of textile manufacturing and the iron industry. To be sure before the 1850s, America

⁶A similar but more complex argument could be made with regards to the growth rates. A sensible assumption is that there is a positive relationship between growth rates and dispersion. A model delivering this relationship would feature vintage capital that is retired at a slower rate during expansions. This would suggest as in the level case that the U.S. at this time would have lower dispersion since China and India are growing more rapidly currently.

was still mainly an economy of small artisanal shops and agriculture. By the close of the century, massive manufacturing firms were beginning to dominate due to increasing returns to scale in manufacturing (O'Brien, 1988). In 1850 14.5% of the population was in manufacturing and nearly 20% by 1880. This came at the cost of the share of the labor force devoted to agriculture which fell by an almost identical amount. Unsurprisingly, this period of U.S. history saw rapid urbanization as people left agriculture and moved to the cities in search of manufacturing and services work. When Abraham Lincoln took office, approximately 16% of the population was urbanized. This percentage had more than doubled by 1900 rising to nearly 40%.

A similar process of industrialization has been playing out again in India and China. Growth in industrial production in China during the first decade of the 21st century averaged over 9%. This growth in manufacturing has also pulled masses of people into the city. By 2008, 46% of the population was urbanized. This was a large rise from 37% during the mid 1980s (OECD, 2005). For India, the process has been a bit slower in the making. Until the early 1990s, India suffered from the so-called Hindu rate of growth of around 3%. This translated into 1.3% rate of per capita growth. This finally began to change in the early 1990s with an acceleration in growth. India has trailed China the whole way through this process. For example, 52% of India's labor force is in agriculture while only 14% is in manufacturing (OECD, 2010). Moreover, fewer people in India have moved to cities with an urbanization rate in 2001 of 28%.

3.2 General Economic Environment

While the economic trajectories of these countries appear quite similar, the differences in regulatory regimes are rather large. Simply put, the U.S. in the 19th century was free of any major economic interventions by the government and, in a positive sense, vigorously defended property rights. To be sure, as Hughes (1978) argues, even going back to colonial times, the U.S. was never totally the libertarian utopia much rhapsodized by some. Still the idea of anti-trust regulation or limiting the size of firms would have been absurd at the time. Things like state-owned enterprises (SOEs) that still account for 40% of employment in China simply did not exist in the U.S. at this time.⁷ Similarly, the level of red tape and bureaucracy imposed during the License Raj in India is unfathomable in

⁷These SOEs can potentially have impacts well beyond lower TFP. A recent paper by Zang et al. (2010) has constructed a growth model with SOEs that shows how China's large foreign surplus could result from the misallocation of credit.

the U.S. today and even more so in the 19th century.

To be sure, there were some other dimensions not strictly related to economic policy that left something to be desired in the U.S. at this time. A major impediment here would have been areas especially on the frontier that lacked law and order. Another possible negative at this time was corruption of public officials from police officers in Chicago to judges on the frontier. For both law and order and the level of corruption, it is difficult to make a meaningful comparison to the India and China of today. The restive areas of China and India are well known as well as various cases of corruption. Law and order is actually one area where it can safely be asserted that the situation in the U.S. today is superior to that of the 19th century.

This hands off approach was both a local and federal policy. Federal spending as a total fraction of GDP was never more than a few percent except during the Civil War when the share of spending reached 17%.⁸ In peacetime, total federal spending averaged around 5%.⁹ For China, total government expenditure inclusive of local spending averaged around 25% of GDP over the HK time period. The major role of SOEs in China's economy is well known. While Zang et al. (2010) comment on how rapidly the private sector has developed, still in 2009 almost 40% of employment is in SOEs, and the private sector did not become the majority contributor to aggregate value added until 2001. The SOEs span the range of industries from banking to oil production and play major roles in all of them.

India led the world in distortionary economic policy for many years stretching back to the country's first prime minister, Jawaharlal Nehru, who was a strong supporter of state intervention. He also was a strong believe that India should be self-sufficient in producing everything it needed. Nearly all economic transactions were planned by the state, and high tariffs were put in place that effectively closed the Indian market to any foreign trade (OECD, 2010). And with that, SOEs filled the gap commanding the vast majority of resources in the economy. This was the part of the system that came to be known as the License Raj. Resources were showered on very capital-intensive heavy industry as well as manual, low-skill industries. There were some small tentative reforms in the 1980s but the major reforms did not begin until the early 1990s in the middle of the HK time period. These included the opening up of the economy to trade and the loosening of

⁸Compare this to the 45% during WWII and the Civil War seems like nothing more than an "unpleasantness."

⁹We do not have good data for local spending, but it is reasonable to assume that this was not large either.

regulations on production. Even though much has changed, still in 2007 40% of value added was produced by SOEs, and the government is completely dominant in certain industries like electricity (OECD, 2010).

Another important policy difference between the U.S. and China involves property rights. For example, it was not until as recently as 2004 that China outlawed arbitrary seizure of private property by the government. Property has been nearly sacrosanct throughout American history.¹⁰ Property rights issues also limit migration in China because peasants choosing to move to big cities forfeit their leases on land with no opportunity for compensation. What has received more attention is China's protection of specifically foreign intellectual property, which has been especially poor. Still many local Chinese entrepreneurs worry about counterfeiting of local products as well. The U.S. has had relatively strong protection of intellectual property rights with a robust patent system enshrined in the Constitution. On the balance when it comes to general economic policy, HK are surely right to use the U.S. as the baseline case of benign institutions case in comparison to China and India. This comparison is only strengthened when considering the 19th century U.S.

3.3 Financial Markets

This period of time in U.S. economic history is one of rapid financial development partly in response to development supported by benign policy. While massive interest rate differentials continued to exist between regions even in 1870, this differential fell to about 1% by 1880 (Davis, 1965; Bodenhorn and Rockoff, 1992). The 1850s witnessed the development of bank clearinghouse to clear notes without the burden of actually shipping specie bank and forth between banks. The clearing houses not only rationalized the payment service, it served as a sort of de facto central bank extending credit to far flung banks when demand for specie was unusually high. This era saw the spread of the commercial paper market, the growth in the stock market, and rise of investment banks. It should be emphasized that the government played no major role in nurturing these institutions and did not regulate their subsequent behavior.

Of course, these institutions are mainly useful for very large firms rather than small businesses, which were then and still are today funded predominantly through local banks. The changes in

¹⁰Whether these rights are strong because of legal protections or rather a convention is a subtle question. See Lamoreaux (2010) for an interesting discussion here.

the local banking market were also salutary. Attack and Passell (1992) note that in 1860 there was approximately one bank for every 10,000 inhabitants. This fell during the Civil War to one per every 20,500 inhabitants, but, by 1890, there was one bank for every 7,700 residents. These new banks were not run by the government. Rather they were private banks operating under first state charters, then free banking statutes, and finally national bank charters. With this freedom and little oversight, the banking sector was prone to experience crises such as the panic in 1907.¹¹

There were some other major policy changes during this era including the arrival of fiat money during the Civil War and the National Bank Act of 1861 and 1864. In order to fund the war effort, Congress issued non-interest bearing Treasury notes that were legal tender for all debts private and public. The supply of these notes was rapidly expanded by Congress effectively doubling the nation's money supply (Attack and Passell, 1992). The resultant depreciation in the value of these so-called greenbacks pushed gold out of actual circulation. It is hard to think that this sort of policy change would be helpful in the efficient allocation of resources.

On the other hand, the National Bank Act was essential in finally establishing a uniform national currency. The legislation effectively taxed private bank notes out of existence and led to the rise of demand deposits. It also provided the first national banking regulatory body, the Office of the Comptroller of the Currency, whose job was to examine banks chartered under this law. The law also allowed like the free banking statutes at the state level any group of individuals to obtain a charter to operate a bank. So while distorting on certain dimensions like the tax on note issuance, the law continued the policy of liberal bank entry and helped put in place the institutions necessary for national monetary and regulatory policy.

The governments of China and India still massively intervene in directing credit. India started from a position during the height of the License Raj where 90% of loans were directed by the government. Interest rates were directly administered and loans over a certain size required the approval of the Reserve Bank of India. Public sector banks still hold something like 75% of total assets to this day. Moreover, banks can only allocate 40% of their total assets freely (OECD, 2010). The Chinese banking sector is still dominated by state-owned banks that have been moving in the direction of greater freedom from political interference.

¹¹The one notable restriction was on branch banking, which presumably would have a larger impact on moderating the response to shocks than the efficient allocation of capital.

China's experience during the current crisis offers a useful example. The government did not rely on an independent central bank to slash interest rates or wait for the market to correct itself. Instead it simply directed its banks to fund a huge expansion of loans. In both countries, other funding sources such as equity markets are still underdeveloped partially due to restrictions on what financial activities are allowed. In summary, the U.S. at this time was much more *laissez faire* in terms of banking regulation. This might have contributed to an exacerbation of business cycle fluctuations, but it was surely more conducive to the efficient allocation of capital.

3.4 Labor Markets

Labor markets in the U.S. were unfettered relative to China and India. Of course in the U.S. during the Civil War, this freedom of movement was somewhat crimped. For the majority of the time period I study, it was as simple as it is today to pack up and move to new opportunities with no prior approval from the U.S. government. In support of these unfettered markets, Ferrie (2005) has found large amounts of migration during this time in U.S. history. To be sure, others such as Wright (1979, 1981) have suggested otherwise in effect arguing that the North and the South were separate labor markets. Moreover, the work by Wright argues that mobility *within* the South was somewhat limited as well even for whites.

Even temporary moves in China require a number of permits from various local authorities. Internal migration is limited not only through explicit restrictions but through other policies. As discussed above, migrants give up their leases to their land with effectively no compensation. They also can lose access to social services like health care and education if they attempt to migrate. (OECD, 2005) These policies result in a massive wage differential between rural and urban areas in China of 3.4 times in 2008 (Candalaria et al., 2009).¹² Indirect evidence for these restrictions come from Au and Henderson (2006) who show that based on estimated agglomeration benefits that many Chinese cities are much too small.

Even though India has no explicit restrictions on internal migration, there is very little of it to speak of. What seems plausible is that the low level of migration is a spillover from very restrictive labor laws. India finds itself only trailing Portugal for the most restrictive regulations. These laws dampen hiring and firing and hence the necessary turnover in the labor market that provides the

¹²it is even larger (about 5 times) if you compare urban areas on the coast to rural ones.

pool from which migrants are drawn. In fact, these type of restrictions are precisely those that would have detrimental effects on TFP in the model of Lagos (2006). In comparison, China is around the OECD average and the modern U.S. has the freest labor markets.

3.5 International Trade

This is one policy dimension where the countries in question are relatively similar. For the U.S., the federal government did use tariffs as a way to raise funds during the Civil War and the tariffs on European goods continued after the war. However, these tariffs did not seem to hurt the openness that much. For example, North (1966) has noted the rapid advance in international trade even in the presence of the tariffs. From 1840 to 1860, both imports and exports nearly tripled. While protecting infant industries was a relevant consideration, Congress passed relatively broad based tariffs rather than ones focused on excluding a particular good. This limited the degree of relative distortions within the U.S.

The incredible growth in the openness of China starts even before 2001 when they joined the WTO. From 2000 to 2005, China's average annual growth rates of export and imports were 25% and 24% respectively. Besides reductions in tariffs, China has established any number of special economic zones where foreign firms can investment and operate in very favorable business climates. This been a major boon to foreign direct investment. To be sure, many foreign companies remain hesitant to jump all the way in due to weak intellectual property rights, a sometimes prickly political leadership, and the undervaluation of the yuan that acts as a broad-based tariff.

India has had a tremendous increase in openness as well. However, much of this come after the time period that HK study. In the early 1990s, India was only slowly beginning to open up the economy from being totally closed to international trade. In the early 90s, tariff revenue totaled 60% of import value and by the 1994, this had fallen dramatically but still stood at 30% in 2005 (OECD, 2010). As for FDI, India remains skeptical with severe limitations on what foreigners can invest in including restrictions on much of the services sector. On the balance, the 19th century U.S. was more open to the world than India during the HK period and probably slightly less open than China, but the differences here are probably not large.

4 Data

The data I use is a sample from the Census of Manufactures described in Attack and Bateman (1999). The first formal COM was conducted in 1840, but scholars have been very skeptical about the quality. Data appear to be coded inconsistently across industries, which makes comparisons almost impossible. Only starting in 1850 is a consistent and trustworthy COM undertaken by the U.S. Quoting Attack and Bateman (1999)

[N]o single early economic data source surpasses the nineteenth-century U.S. federal census manuscripts in quality, in consistency, or in comprehensiveness; from mid-century onward, the census enumerations offer a unique historical record detailing the transformation of the United States from an agricultural to an industrial economy.

The strengths of the data are quite clear. It contains a rich set of variables including data on the wage bill, revenue broken down by price and quantity,¹³ capital stock,¹⁴ and inputs. It also has information regarding the type of power used such as steam or animal and with this, the total horsepower employed. There are other questions that are relevant for a bygone era such as number of children employed and whether the factory is seasonal. Industries are 3-digit SIC codes where Attack and Bateman (1999) (AB) could identify them. This is a slightly higher level of aggregation than HK who use 4-digit codes. In some cases, AB could only place firms at the 2-digit level or in a conglomerate category that spans multiple industries.¹⁵ Firms above \$500 in output were to be enumerated. AD point out that in 1870, enumerators were “explicitly cautioned” to not omit the smallest shops. The electronically available sample only contain representative samples of about 5000 observations.

The Indian data used by HK is a census of all large manufacturing plants and a random one-third sample of smaller plants. HK apply the provided sampling weights to create a representative sample. The Chinese data is to be precise about *firms* not plants and is a complete census of firms with revenue greater than 5 million yuan (about \$600,000). Both of these datasets do not include any of the small scale production that goes on in informally in these two countries. Presumably,

¹³This breakdown by price and quantity is even lacking in modern manufacturing data such as the Annual Survey of Manufactures.

¹⁴This variable is not constructed using something like the perpetual inventory method. The capital stock measures in the Indian and Chinese data are book values.

¹⁵I have experimented with excluding these conglomerate firms with little effect on the results.

any differences in enumeration of smaller units will translate into the measure of dispersion with data that covers these firms better showing how measures of dispersion.

In addition to the limited sample size, the data does not allow for uniquely identifying firms over time. The names were simply not recorded by Attack and Bateman due to memory constraints.¹⁶ So I cannot attempt some of the exercises that HK try such as looking at the relationship between age and TFPR or future input growth and TFPR.¹⁷

5 Main Findings

Using the mid 19th century U.S. data, the results of the HK calculations for TFPR and TFPQ are presented in Tables 3 and 2, respectively. There does not appear to be a clear trend in these measures of dispersion over time. While there does not appear to be a clear break due to the Civil War between 1860 and 1870 in the dispersion of TFPQ, there does appear to be a slight jump in TFPR. This suggests that the Civil War may have distorted resource flows across the U.S., which seems entirely reasonable. However, this would require much more evidence to validate this claim.

In Tables 5 and 4, I report the average over the 3 censuses for China and India from HK versus the average over the 19th century U.S. censuses. The statistics for the TFPR for the 19th century U.S. economy are nearly identical to those of China and India for a variety of dispersion measures. All of these measures are also appreciably lower than the modern U.S. data. Figures 1 and 2 plot the distribution of log TFPR and log TFPQ relative to the log of the industry average for the full sample. For TFPQ, there is also very little difference between the developing countries and the contrast is not quite so stark with the modern U.S.

If it is policy that explains these differences, then what are we to make of these results? We started with a completely different country in terms of policy but a similar level of development in the form of the 19th century U.S. and ended up with very similar levels of misallocation. This casts doubt on the relation between policies and the measures of distortions suggested by HK. Whereas the policy distortions in India and China are well known in terms of state owned enterprises and previously the License Raj in India, it is not so easy to what those explicit barriers are in the

¹⁶They started collecting this data on punchcards!

¹⁷The modern COM data while anonymized is linked across years allowing for people to exploit dynamic patterns in the data.

U.S. at this time. As I argued above, economic institutions were rather conducive to allocating resources. Government spending was limited. There were few market restrictions. Bank decisions were not overseen directly by the government. And where the government did play a larger role in developing communication and transportation networks, its effect appears to have been salutary.

Simply put, there is no good policy reason why the U.S. at this time should have such a distorted capital allocation. The only thing that it shares with China and India of the late 20th century is a similar level of economic development as reflected in real GDP per capita. I interpret these results as suggesting that part of the natural process of development is capital reallocation. Policy surely does play some role here, but I think it is much less than we think while development itself is the main driver. What this means is that HK succeeded along the accounting dimension while overstating the case for the institutional one. Rather than a cause of low income per capita, capital misallocation seems to be an effect.

We can also look at the industry-level for another piece of evidence supporting my interpretation these results. Presumably if reallocation is a process involved naturally in any growth process, then it should be present at the industry level as well. To test this idea, I regress subsequent growth in an industry's value added on the present level of dispersion of TFPR in that industry. If high levels of "misallocation" are really just a signal of a developing industry, then I would expect a positive relationship between these two variables. This is precisely what I find in Table 8. There is a massive effect of this measure of misallocation on growth. A one standard deviation increase in the degree of misallocation increases the growth rate of value by 25 percentage points. A similarly sized effect is present for capital as well. The results are similar if we instead use the interquartile range as a measure of dispersion. A similar effect is present though not statistically significant if we instead use the growth rate of the average value added in an industry.

It is easy to find examples as well of industries with highly distorted capital allocations that retrospectively we can identify as the progressive ones. For example in 1850, the coke, iron and minerals industries are some of the most highly distorted industries, and we would think of these as constituting part of the technological frontier at the time. To be sure, one can find other industries that appear highly distorted but do not appear to be "high-tech." For example, the salt industry, which would not be considered particularly novel now or then, is consistently near the top in terms of misallocation. On the balance though, the supportive examples outweigh the negative ones

providing evidence for the idea that misallocation and reallocation are simply parts of the process of development and especially infant industries.

In addition, it does not appear that the results are driven by a higher degree of measurement error in the old data. In Table 7, I run the HK checks for measurement error using the difference between the modern U.S. regression coefficient of revenue on cost and the one estimated here as an estimate of the degree of measurement error. If there is no measurement error (at least in these two variables), then the coefficients should be equal. The extent to which they differ provides a measure of the degree of measurement error. What is most important to note is that the coefficients for my data do not look much different from the ones for China and India. Hence, it does not appear to be the case that the degree of dispersion driven by measurement error in the U.S. in the 19th century is much different than the developing countries. Following HK, the estimates imply that potentially a few percent of the difference in the level of dispersion between the modern U.S. and its 19th century counterpart can be explained by a higher degree of measurement error in the older sample.

I also have repeated the analysis dropping industries with fewer than 100 observations and excluding observations from Confederate states. One may worry that industries with very few observations are simply too noisy due to small sample sizes. One may also worry that data in the Confederate states may simply be of our lower quality due to difficulties in enumerating these states at this time. Table 6 shows the results for these subsamples. As is quite obvious, the measures of dispersion are nearly identical across these subsamples. Figures 3 and 4 are the distributions when we exclude industries with few observations. As is evident, there is little difference here as well.

6 Implications for the Development of the American Economy

I now turn to a calculation of what fraction of manufacturing TFP growth over the last century of U.S. history can be attributed to a better within industry allocation of resources. Work by Petrin et al. (2009) has shown that for the U.S. between 1976 and 1996, reallocation has contributed between 1.6 and 2% to the growth rate in TFP while technical efficiency itself has only grown .2 to .6% per year. Moreover, while technical efficiency is highly volatile and procyclical, reallocation rarely detracts from overall productivity growth.

HK envisioned a hypothetical experiment whereby the distribution of wedges for China and India were set to that of the U.S. Given this more efficient allocation of resources, they calculated how much higher China and India's TFP would be. Depending on parameters, in particular the elasticity of substitution σ , they could explain 50% of the gap between manufacturing TFP in China as compared to the U.S. and 35% in India. They concluded that capital misallocation could possibly explain a sizable fraction of differences in average levels of income.

To do the calculation, I first need to specify how manufacturing final output Y is produced from sectoral-level output Y_s . I again follow HK and assume a simple Cobb-Douglas specification

$$Y = \prod_s Y_s^{\theta_s} \quad (11)$$

where $\sum_s \theta_s = 1$. Now we can calculate what TFP would be in the efficient allocation.¹⁸ In the efficient allocation, TFPR will be equalized across firms in a given sector. HK show that the gain in output can be calculated as

$$\frac{Y}{Y_{eff}} = \prod_{s=1}^S \left[\sum_{i=1}^{M_s} \left(\frac{A_{si} \overline{TFPR}_s}{A_s \overline{TFPR}_{si}} \right)^{\sigma-1} \right]^{\frac{\theta_s}{\sigma-1}} \quad (12)$$

where Y_{eff} is output in the efficient allocation. This exercise fixes the total amount of inputs and calculates how much output could be increased by reallocating resources between firms. This is why the increase in output from the actual to the efficient allocation is equal to the potential gain in TFP. Because of data limitations, it is not possible to operationalize this calculation in 1880. So these calculations will be restricted to the first three censuses.

Table 9 displays these calculations. I follow HK in choosing the U.S. in 1997 as the baseline. Even for this modern time and place, HK estimate that there is potential for a 43% improvement in TFP. Still there are even more massive potential gains from a better allocation of resources in the 19th century. For 1860, TFP could almost be doubled while even in 1870 the lowest year, gains of 60% could be had. Between 1890 and 1997, U.S. manufacturing TFP increased by 465% (Field, 2010). If we take 1870 then, this implies that approximately 14% of the total gain in TFP over the last 100 years can be attributed to a better allocation of resources. The other years imply even

¹⁸Note that θ_s can be calculated simply from the share of total output from sector s .

higher fractions of up to 38%. The key parameter driving these results is the degree of substitution between goods, σ . HK choose the most conservative value from the modern literature, and I follow them. Choosing a value of σ at the high end would imply orders of magnitude greater gains from reallocation.

I fully admit the heroic nature of this calculation. The calculation here leaves no room for measurement error (though the data does not appear to suffer any greater amounts of measurement error than the Chinese or Indian data) or model misspecification. As HK point out, it is easily the case that either of these issues could lead to an overstatement of the efficiency gains. Possible overstatement due to measurement error is quite obvious. If much of the variation across plants is noise in the output and input measures, then clearly the true level of dispersion and how much can be gained from reallocation is much lower. Model misspecification in the form of, say, assuming a Cobb-Douglas production function can lead to incorrect inference to the extent that this biases productivity measurements. Assume that there are actually increasing returns to scale, then productivity measures for larger plants would tend to be upward biased again leading to an overestimate of the gains from reallocation. And of course, there is the question of applying this method over long sweeps of history much like many of the reservations that surrounded the use of Solow residuals.

In some ways though, this number should really be thought of as a lower bound on the importance of reallocation. The reason is that the HK methodology only considers misallocation *within* a sector as compared to misallocation *between* sectors. If we think about the massive shifts over the course of history, it has been between sectors from agriculture to manufacturing and now onto services. This sort of movement goes completely unaccounted in this setup for two reasons. The most obvious reason is that the data employed is restricted to manufacturing. However, even with an extended dataset with other sectors like agriculture, the setup rules out frictions in reallocating between sectors. The final good producer faces no “wedges” that limit substitution between different sectors. The setup also rules out by assumption reallocation between entering and exiting firms by assuming a static set of firms in a given industry. These shortcomings make these calculations even more striking.

7 Conclusion

This paper contributes to both the historical literature on the development of the American economy and more broadly to economists' thinking on development. The results suggest that a sizable fraction of subsequent manufacturing TFP growth in the 20th century can be tied to a better allocation of resources. For the latter, this paper shows that the mapping from economic policy to the allocation of capital is not straightforward. It appears instead that part of the process of development itself is a natural reallocation of capital towards more efficient ends, totally independent of policy decisions. And this endogenous process is now what needs a theory. Future work should attempt to fill in the gap between the results presented here for 19th century America and those in HK for the 20th century. In the end, much remains to be learned about development through the lens of economic history.

References

- Acemoglu, D., S. Johnson, and J. Robinson (2001). The colonial origins of comparative development. *American Economic Review* 91, 1369–1401.
- Acemoglu, D., S. Johnson, and J. Robinson (2005). Institutions as the fundamental cause of long-run growth. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*. Elsevier.
- Acemoglu, D. and J. Robinson (2009). *Economic origins of dictatorship and democracy*. Cambridge University Press.
- Atack, J. and F. Bateman (1999). Nineteenth-century U.S. industrial development through the eyes of the Census of Manufactures: A new resource for historical research. *Historical Methods* 32, 177–188.
- Atack, J., F. Bateman, and R. Margo (2004a). Capital deepening in United States manufacturing, 1850-1880. Mimeo, Vanderbilt University.
- Atack, J., F. Bateman, and R. Margo (2004b). Skill intensity and rising wage dispersion in Nineteenth-century American manufacturing. *Journal of Economic History* 64, 172–192.
- Atack, J., M. Haines, and R. Margo (2008). Railroads and the rise of the factory: Evidence for the United States, 1850-1850. NBER Working Paper, 14410.
- Atack, J. and P. Passell (1992). *A new view of American Economic History*. Norton.
- Au, C.-C. and J. Henderson (2006). Are Chinese cities too small? *Review of Economic Studies* 73, 549–576.
- Banerjee, A. and K. Munshi (2004). How efficiently is capital allocated? Evidence from the knitted garment industry in Tiripur. *Review of Economic Studies* 71, 19–42.
- Bodenhorn, H. and H. Rockoff (1992). Regional interest rates in antebellum America. In C. Goldin and H. Rockoff (Eds.), *Strategic factors in Nineteenth century American economic history*. University of Chicago Press.
- Candalaria, C., M. Daly, and G. Hale (2009). Beyond Kuznets: Persistent regional inequality in China. Unpublished, Federal Reserve Bank of San Francisco.
- Chari, V., P. Kehoe, and E. McGratten (2007). Business cycle accounting. *Econometrica* 75, 781–836.
- Davis, L. (1965). The investment market, 1870-1914: The evolution of a national market. *Journal of Economic History* 25, 355–393.
- Ferrie, J. (2005). The end of American Exceptionalism? Mobility in the U.S since 1850. *Journal of Economic Perspectives* 19, 199–215.
- Field, A. (2010). *A Great Leap Forward: 1930s Depression and U.S. Economic Growth*. Yale University Press.
- Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review* 98, 394–425.

- Guner, N., G. Venture, and Y. Xu (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics* 11, 721–744.
- Hall, R. (1997). Macroeconomic fluctuations and the allocation of time. *Journal of Labor Economics* 15, S223–S250.
- Hall, R. and C. 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. and P. Klenow (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics* 124, 1403–1448.
- Hsieh, C.-T. and P. Klenow (2010). Development accounting. *AER: Macroeconomics* 2, 207–223.
- Hughes, J. (1978). *The Governmental Habit*. Basic Books.
- Hutchinson, W. and R. Margo (2006). The impact of the Civil War on capital intensity and labor productivity in southern manufacturing. *Explorations in Economic History* 43, 689–704.
- Lagos, R. (2006). A model of TFP. *Review of Economic Studies* 73, 983–1007.
- Lamoreaux, N. (2010). The mystery of property rights. Unpublished, Yale University.
- Midrigan, V. and D. Yu (2009). Finance and misallocation: Evidence from plant-level data. Unpublished NYU.
- North, D. (1966). *The economic growth of the United States, 1790-1860*. Norton.
- North, D. (1989). *Institutions and economic growth: An historical introduction*. Elsevier.
- O’Brien, A. (1988). Factory size, economies of scale, and the Great Merger Wave of 1898-1902. *Journal of Economic History* 48, 639–650.
- OECD (2005). OECD Economic Surveys: China.
- OECD (2010). OECD Economic Surveys: India.
- Parente, S. and E. Prescott (2000). *Barriers to riches*. MIT Press.
- Petrin, A., T. White, and J. Reiter (2009). The impact of plant-level resource reallocation and technical progress on U.S. macroeconomic growth. Mimeo, University of Minnesota.
- Restuccia, D. and R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics* 11, 707–720.
- Wright, G. (1979). Cheap labor and southern textiles before 1880. *Journal of Economic History* 39, 655–680.
- Wright, G. (1981). Cheap labor and southern textiles, 1880-1930. *Quarterly Journal of Economics* 95, 605–629.
- Zang, F., K. Storesletten, and F. Zilibotti (2010). Growing like China. Mimeo, University of Zurich.

	GDP per capita	GDP Growth	Growth per capita	% Agriculture	% Urban
China 1998-2005	4407	8.96	8.17	53	32
India 1987-1994	1349	5.38	3.32	67	25
U.S. 1850-1880	2925	4.73	2.16	44	26

Table 1: Some summary statistics for the three economics. For China and India, data are from IMF World Economic Indicators. For U.S., real GDP per capita in 2005 U.S. dollar is from measuringworth.com. This is translated into PPP dollars and is unit reported for China and India. For China, agricultural fraction is through year 2000 while for India, it is through 1990. This data is from the World Development Indicators. % agriculture for the U.S. and urban is for 1880 only.

	1850	1860	1870	1880	Average
S.D.	.65	.71	.79	.62	.71
75%-25%	.78	.92	.97	.77	.86
90%-10%	1.57	1.76	1.92	1.51	1.73
<i>N</i>	5174	5070	3769	3326	17339

Table 2: Dispersion of log TFPR scaled by industry TFPR from 19th century U.S. Census of Manufactures.

	1850	1860	1870	1880	Average
S.D.	.91	1.03	1.01	.97	.98
75%-25%	1.45	1.81	1.57	1.56	1.6
90%-10%	2.31	2.60	2.57	2.47	2.58
<i>N</i>	5174	5070	3769	3326	17339

Table 3: Dispersion of log TFPQ scaled by industry TFPQ from 19th century U.S. Census of Manufactures.

	India 1987-1994	China 1998-2005	U.S. 1850-1880	U.S. 1977-1997
S.D.	.68	.67	.71	.45
75%-25%	.80	.87	.86	.47
90%-10%	1.65	1.69	1.73	1.08

Table 4: Comparison of Dispersion for log TFPR scaled by industry TFPR

	India 1987-1994	China 1998-2005	U.S. 1850-1880	U.S. 1977-1997
S.D.	1.19	.99	.98	.83
75%-25%	1.56	1.33	1.6	1.16
90%-10%	3.04	2.53	2.58	2.15

Table 5: Comparison of Dispersion for log TFPQ scaled by industry TFPQ

	Full sample	Exclude industries with few observations	Exclude Confederate States
S.D.	.71	.70	.70
75%-25%	.86	.84	.83
90%-10%	1.73	1.71	1.71

Table 6: Comparison of Dispersion for log TFPR scaled by industry TFPR for a variety of sub-samples of the Attack-Bateman sample.

	U.S. 1850-1880	China	India	U.S. 1977-1997
Inputs on Revenue	1.02 (.002)	.98	.96	1.01
Revenue on Inputs	.93 (.002)	.82	.90	.82

Table 7: Regression of inputs on revenue and revenue on inputs following HK. Number in parenthesis is standard error for regressions using 19th century data.

	Growth of value added	Growth of capital
Standard deviation of TFPR	2.41 (1.24)	2.47 (1.14)

Table 8: Regression of future growth of an industry on current level of misallocation. These regressions include time fixed effects and standard errors in parentheses are clustered by industry.

	1850	1860	1870
% TFP Gain	125	175	61

Table 9: TFP gains from equalizing TFPR relative to 1997 U.S. benchmark. These numbers are calculated as the ratio between Y_{Eff}/Y to the U.S. ratio in 1997 taken from HK (1.439) subtracted 1, and multiplied by 100.

Figure 1: Smoothed histogram of the log of scaled TFPR. See text for description of estimation methods.

Figure 2: Smoothed histogram of the log of scaled TFPQ. See text for description of estimation methods.

Figure 3: Smoothed histogram of the log of scaled TFPR for a subset of industries. See text for a list of industries included.

Figure 4: Smoothed histogram of the log of scaled TFPQ for a subset of industries. See text for a list of industries included.

Figure 5: Smoothed histogram of the log of scaled TFPR excluding Civil War ravaged states.

Figure 6: Distribution of the log of scaled TFPQ excluding Civil War ravaged states.