This paper was prepared by Professor Chang-Tai Hsieh, University of Chicago, for the “Global Dialogue on the future of productivity growth: Towards an OECD productivity network”, Mexico City, 6-7 July 2015.

This paper should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments are those of the author.
When we talk about GDP growth, or investment rates, or export growth, we sometimes slip into language that gives the impression that these are instruments rather than the outcome of a complex process that we are only beginning to understand. What are ultimately behind this complex process are people and firms. Growth, investment, and exports are the outcome of the processes by which people with ideas start firms; some of them turn out to be successful because they are able to deliver products for which global customers are willing to part with their hard earned money. When this happens systematically for a large number of firms, growth takes off. This, in a proximate sense at least, is what has happened in some countries, such as China, but has generally failed to occur in many countries around the world.

This perspective on growth raises three important questions. First, where does the productive capacity of firms come from? Second, what are the barriers that make it difficult for productive firms to get access to the resources they need, and that allow less successful firms to dominate the market. Third, what are the constraints that prevent individuals with the potential to create world class firms from doing so in the first class?

There has been a great deal of exciting research on these questions in the last few years. The starting point of this research is the well-established fact that a large proximate determinant of differences in output per worker across countries is the residual term or total factor productivity (TFP). Similarly, the bulk of the cross-country differences in growth rates are driven by differences in the growth rate of aggregate TFP. Hall and Jones (1998), Caselli (2005), and Hsieh and Klenow (2009) are useful references on this evidence. Although many have interpreted this fact as evidence of differences in technology playing a large role, recent research
by Diego Comin shows that technology diffusion, at least in the modern era, is relatively quick. What this literature does is to take seriously the notion that aggregate TFP may not be driven by technology but rather by allocative efficiency. In the Soviet Union, or in China under Mao, we don’t think that low TFP was driven by low technology but rather by inefficiencies in the allocation of resources between firms.

To understand the key ideas behind this literature, it is useful to invest in some notation. The building block of this literature is the basic idea that firms differ, and we don’t necessarily want all the resources to be allocated to one firm. For example, suppose that there a number of firms in a country and the output of each firm is given by:

$$Y_i = A_i F(K_i, L_i)$$

If each firm produces different products, we don’t necessarily want all the inputs allocated to the firm with the highest $A_i$, as we value having access to a variety of differentiated products. Instead, what we want is for resources to be allocated across firms to equalize the revenue productivity of the firm, or $P_iA_i$. Resources are misallocated when revenue productivity differs between firms. And reallocation increases aggregate TFP and generates growth when resources flow to firms with high revenue productivity and away from firms with low revenue productivity.

This basic idea extends to any framework that allows for heterogeneous firms to exist in equilibrium. For example, suppose that firms make identical products but what supports an equilibrium where firms differ in productivity is that each firm faces a “span of control.” This can be because of limited managerial capacity or because of supply constraints. In this case, aggregate efficiency is maximized when resources are allocated across firms to equalize average revenue productivity per inputs (revenue per weighted average of capital and labor in a model where the inputs are capital and labor). Resources are misallocated when average revenue productivity differs between firms, and reallocation is growth enhancing when resources flow to firms with high average revenue productivity.
Using detailed micro-data from the manufacturing censuses in India, China, and the US, Hsieh and Klenow (2009b) show that there are substantial gaps between firms in revenue productivity in India and China. These gaps are also present in US data, but to a much smaller extent. In India in 1994 and China in 1998 for example, revenue productivity of firms in the 90th percentile differed from that of firms in the 10th percentile by a factor of five. In the US in 1997, the gap in revenue productivity between firms in the 90th percentile and firms in the 10th percentile differ by a factor of 3. These gaps in revenue productivity translate into substantial gaps in aggregate TFP. Hsieh and Klenow (2009b) calculate that aggregate TFP would increase by 43% in the US, by 115% in China in 1998, and by 127% in India in 1994 if resources were to be reallocated between firms such that the revenue productivity is equalized across firms.

We now have a large body of evidence of the gaps of revenue productivity at the micro-economic level, largely thanks to the detailed firm level data available for a growing number of countries. A project spearheaded by Santiago Levy at the IDB (IDB 2010) provides detailed evidence on these gaps for a large number of countries in Latin America. Bartelsman, Haltiwanger, and Scarpetta (2011) provide similar evidence from micro-economic data for a number of countries in Europe. This body of evidence indicates that the gaps in revenue productivity are wide, consistent with misallocation.

It is worth discussing at this stage the widely used methodology developed by Olley-Pakes to measure resource misallocation. This methodology is used by the Bartelsman, Haltiwanger, and Scarpetta (2014) paper mentioned earlier. Andrews and Cingano (2014) is a recent paper that applies this methodology to the OECD countries. The issue with the Olley-Pakes methodology, as stated clearly by Bartelsman, Haltiwanger, and Scarpetta, is that in any model with firm heterogeneity, allocative efficiency implies that revenue per unit of \textbf{variable} inputs should be equated across firms. Put differently, the covariance between revenue per unit of variable inputs and firm size should be zero when resources are allocated efficiently. A positive covariance implies allocative inefficiency, in the same way that a negative covariance implies inefficiency. It is possible to argue that a higher covariance between measured revenue productivity and firm size implies efficiency if revenue productivity also includes fixed inputs in the denominator. This is, for example, the logic in Bartelsman, Haltiwanger, and Scarpetta, but then the question is
whether the differences in the covariance in the measured revenue productivity and firm size across countries (e.g. in the Andrews and Cingano paper mentioned above) reflects differences in fixed costs or something else.

The literature has largely focused on measuring the static effect of these forces, but these forces are likely to also have important dynamic effects. Specifically if more efficient establishments face bigger distortions, then these undermine the firm’s incentive to invest in better technology. Put differently, there are two effects of resource misallocation. First, there is the static effect. Second, there is the dynamic effect of resource misallocation on the growth in firm productivity. Parente and Prescott (2000) and Schmitz (2005) has highlighted this effect in several case studies. Hsieh and Klenow (2014 present evidence from firm level censuses in India and Mexico consistent with this mechanism. They show that while US firms grow by a factor of eight by the age of forty, Mexican firms only double in size, while Indian firms do not appear to grow. This evidence is for India, Mexico and the US, but there are also interesting differences between OECD countries. Criscuolo et al (2014) and Bravo-Biosca et al. (2013) document the differences in the life cycle of firm employment between OECD countries.

The challenge then is to identify the precise policies and institutions behind these gaps in the micro-economic data. Here the potential list is large, and I will discuss six forces that might be important and for which we have some evidence. Identifying the precise forces is obviously critical if we are to guide policy.

First, we have a large body of evidence that there can be substantial costs to adjusting labor and capital inputs. Caballero, Engel, and Haltiwanger (1997) provide evidence on labor adjustment costs and Cooper and Haltiwanger (2006) on capital adjustment cost. What we still have limited evidence on is whether adjustment costs might be responsible for the gaps in revenue productivity observed across firms between rich and poor countries. Haltiwanger, Scarpetta, and Schweiger (2008) show that there is less labor reallocation in countries with more rigid labor regulations. Asker, Collard-Wexler, and de Loecker (2014) show that the dispersion in capital productivity is correlated with productivity volatility in a small sample of countries, although the underlying sources of productivity volatility are unexplained. In work in progress with Pete
Klenow, I show that there is much less volatility in employment in Indian manufacturing compared to the US, almost as if firms face large costs to adjusting employment. This is consistent with the evidence in Aghion, Burgess, Redding and Zilibotti (2008) who focus on the effects on rigid licensing laws in India. Andrews, Criscuolo and Menon (2014) focus on differences across the OECD countries in the ability of innovative firms to attract resources that would enable them to grow.

Second, another source of dispersion in revenue productivity is ownership of firms by the state or political cronies. We have some evidence on the effect of state ownership on the dispersion in revenue productivity. Hsieh and Klenow (2009b) shows that state owned firms in China had substantially lower revenue productivity in the late 1990s but the gap narrowed after that due to the closure and privatization of many state owned firms. A detailed study of state owned firms in India by McKinsey (2001) shows substantial gaps in labor productivity between state owned firms and privately owned firms in the same sector. In the telecoms sector, labor productivity of private firms is three times higher than that of state owned firms. In the retail banking sector is more than five times higher in private firms compared to state owned firms. To be clear, these gaps do not necessarily indicate that TFP of state owned firms is lower. Although it is likely to be the case that TFP of state owned firms is lower, the evidence of lower labor productivity simply indicates that the marginal product of labor is lower in state owned firms. Here the policy solution is clear, although politically difficult. Privatize the state owned firms, and treat the privatized firms on the same terms as the other firms.

However, even when state owned firms are privatized, it may be the case that these firms are owned by political cronies who use their political power to obtain the same preferences as state owned firms. Hsieh and Song (2015) show that this is what happened to the privatized state owned firms in China. More broadly this raises the point that ownership by politically connected cronies is potentially more important quantitatively, but identifying politically connected firms and measuring the effect of politically connected firms in the data is obviously difficult.

Third, an important feature of many poor countries is the large informal sector. Lagakos (2009), for example, shows that the modern retailing accounts for 67 percent of total employment in the
retail sector in the US. The equivalent numbers are 21 percent for Brazil, 15 percent in El Salvador, 23 percent in Mexico, 15 percent in the Philippines, and 19 percent in Thailand. And what is surprising is that modern retail stores in countries where their employment share is lower than that of informal retailers are also countries where labor productivity of modern retail stores is significantly higher than the informal retail stores. Lagakos (2009) shows that labor productivity of modern retailers is three times higher than that of informal retailers in Brazil, four times higher in El Salvador, 3 ½ times higher in Mexico, six times higher in the Philippines, and four times higher in Thailand.

Hsieh and Klenow (2014) provide similar evidence for the manufacturing sectors in India and Mexico. Informal manufacturing establishments in India, defined as establishments that are not formally registered, account for 80 percent of total manufacturing employment in 2005. In Mexico, almost all manufacturing establishments are formal in the sense of being formally registered. However, if we define informality in Mexico as firms that are not paying social security taxes (either legally by only employing unpaid family workers or illegally by explicitly not paying the social security tax), informal establishment account for 30 percent of total manufacturing employment in 2008. Informal establishments in India and Mexico are also significantly smaller than formal establishments. Figure 1 plots the distribution of establishment size in India and Mexico for informal and formal establishments.

Figure 1: Size of Formal and Informal Establishments in India and Mexico
The typical informal establishment employs four workers in India and about ten workers in Mexico. Formal establishments employ 20 worker in India and about 50 workers in Mexico.

All of this suggests that a key proximate reason why poor countries are poor is because modern formal firms find it difficult to obtain resources and/or capture market share. We still have a very limited understanding of what exactly are the forces behind the prevalence of many informal and unproductive establishments. The InterAmerican Development Bank (2010), and
Santiago Levy (2008) in particular, have argued that the patterns of informality, at least in Latin America, are due to the nature of the tax system and the social protection programs. A glance at the World Bank *Doing Business* indicators suggests that high costs of doing business may also have something to do with it. But the evidence on whether any of these forces are the right ones is still absent.

Although the effect of such programs are likely to be smaller in the OECD countries, a recent paper by Garicano, Le Large, and Van Reenan (2013) looks at the effect of labor laws in France that kick in for firms with more than 50 employees. They find that this labor law has a sizable effect on the French firm size distribution, and the aggregate effect on TFP can be as large as 5% of GDP.

Fourth, deep social forces may also have an important effect on the allocation of talent. These forces may reflect the legacy of gender and race in the US, caste in India, economic and ethnic background in some Latin American countries, or the effect of second generation managers in family firms in many countries. In the US for example, the fraction of white women in high skilled occupations (defined as lawyers, doctors, engineers, scientists, architects, and executives/managers) increased from 6 percent in 1960 to 21 percent by 2008. The share of black men in the high skilled occupations increased from 3 percent in 1960 to 15 percent in 2008. The corresponding share of white men increased much more modestly, from 20 percent in 1960 to 25 percent by 2008. In Hsieh, Hurst, Klenow, and Jones (2012), we show that the convergence in occupations between white men and the other groups can explain 15 percent of the growth in aggregate TFP in the US from 1960 to 2008. Adalet McGowan and Andrews (2015) provide similar evidence for a broad sample of OECD countries, showing that skill mismatch is correlated with differences in aggregate labor productivity.

It is likely that such forces are also present, and perhaps even more important in other countries. And changes in these forces may have played an important role in aggregate growth. Hnatkovska, Lahiri, and Paul (2012) the gap in years of completed education between non-disadvantaged and disadvantaged castes in India was 4.1 years of schooling for women and 2.3 years of schooling for men. The gap is still present today, but has shown a marked decline. In
2004, the schooling gap for women had declined to 2.2 years and for men to 1.7 years. The prevalence of family run firms in many countries also suggests that frictions in occupational choice may be important. Caselli and Gennaioli (2012) and Bloom et al (2012) suggest that first born sons of the firm’s founders are particularly bad managers. Yet, a large number of firms are second generation family firms. This fact suggests that the combination of financial frictions and costs of monitoring non-family members must be important, and one of the implications is the inefficiency in the allocation of talent to top management in many firms.

Fifth, it is likely that internal trade barriers may also play an important role in the efficiency of resource allocation. David Atkin and Dave Donaldson show that internal trade barriers may be very large in poor countries (2012). In the agricultural context, there is some research that suggests that lowering transportation costs in Sierra Leone led to large gains among farmers and also improved competition amongst traders by reducing search costs. In the industrial sector, internal trade costs might have similar effects on firm productivity. It seems likely that if access to input and output markets is critical for modern industrial firms, costs that make it more difficult for firms to access these networks will affect the incentives of firms to invest in firm productivity. If these effects turn out to be large, this will further emphasize the significant gains to be had from further integrating these input and output markets. On this point, recent papers by Bourles et al (2013) and Barone and Cingano (2012) show that internal barriers in service industries in the OECD countries can have important negative effects on firms that use these inputs.

So far, we have discussed firms without taking location into account. Yet, in developed countries, the bulk of economic wealth is generated by people and firms operating within a small number of cities. We do not completely understand why cities are so important, but it seems plausible that cities create business “ecosystems” where firms can obtain the specialized inputs they require, enjoy easier access to their customers, where knowledge and ideas “spill over” across firms, and where skilled workers are located because of their desire for urban amenities and opportunities. Regardless of the underlying reasons, all the evidence indicates that without effective cities, it simply would not be possible to generate sustained growth.
Yet when one looks at a typical city, say in Sub-Saharan Africa or in South Asia, it is hard to escape the view that those cities that are most conducive to firm clustering are also among the world’s most dysfunctional cities. Perhaps the best example of this dysfunction is to look at cities such as Mumbai or Lagos that have experienced large increases in their populations. The population increases in these cities must reflect the increased labor demanded as a result of firm entry or the productivity growth of existing firms. Yet, despite the potential for these businesses to grow, most of them do not. There are many reasons why firms may not grow, but perhaps it’s useful to look at the housing market to illustrate the type of forces operating in cities that might be important.

When one looks at the housing market in such cities, a first order fact is the predominance of slums and informal housing. Slums and informal housing does not imply that the housing is cheap. Although we do not currently have any reliable numbers, it seems likely that when properly measured in quality adjusted terms, the cost of housing in urban areas in poor countries may be significantly more expensive than in many developed countries, certainly relative to per capita income and also, in some cases, in absolute terms as well. Without further measurement, it is not yet clear whether this is the case and this will be a hole in the global knowledge that we will attempt to fill in. Within formal housing markets, where one can observe a market price, some of the most expensive housing in the world is precisely in cities such as Mumbai and Delhi that also have the largest number of slums and informal housing. If the real cost of housing is high, this matters for at least three reasons. First, firms that want to expand will have to offer higher nominal wages to attract the necessary workers or, in extreme cases, may not be able to locate sufficient numbers of workers at all. Second, the cost of housing is probably highly correlated with the cost of land and other resources that businesses need to expand. If it is difficult to acquire office space or new facilities, firms may be forced to remain small. Both of these factors may act as significant disincentives to the expansion required for the firm-level productivity growth discussed above. Again, this is simply a hypothesis, as further research needs to be done in order to ascertain the relevance of this force. Third, rigid housing markets inhibit labor mobility. Hsieh and Moretti (2015) argue that rigid housing markets in US cities such as New York and San Francisco have a sizable effect on labor allocation across cities and ultimately on aggregate productivity growth. Caldera Sanchez and Andrews (2015) provide
related evidence from a broader sample of OECD countries.

Why these housing markets operate in this way is yet to be explained. But it seems plausible that there are four critical ingredients of an efficient housing (and business construction) market that are missing in poor countries. First, there is the absence of a housing finance industry. In the UK, building societies played this role in the 19th century. In Singapore and Hong Kong, this role was taken on by governments that invested in massive urban housing projects. Second, there is the absence of a legal structure that provides titles and legally enforces property rights over land and housing (De Soto 2000; Field 2007). Third, the housing and construction industry appears to be quite inefficient, as recent in Mozambique has pointed out (Marrengula et al. 2012). Finally, local governments simply do not have the capacity to provide the public goods necessary, be these roads, water, sewage, electricity, or other elements that are important complements to housing (World Development Report 2009). Again, the extent to which these forces are important remains to be investigated.

These same forces are likely to affect other public goods that may be just as important for a firm’s growth and productivity. Lack of water, sewage systems, and reliable electricity all have direct effects on a firm’s productivity. In addition, they can negatively affect the establishment of firms producing intermediate goods necessary for growth in the first place. Put simply, can one expect entrepreneurs to set up a semiconductor plant in cities where the power supply goes out for days at a time? How can one attract the skilled workers necessary to build a high-technology industry when the urban amenities are lacking?

This raises difficult questions about the optimal governance structure of local governments. Indeed, it seems plausible that countries where local governments are accountable perform better than countries where local governments are corrupt and where courts are inefficient. But it also seems likely that the quality of local governments are more complex than simple measures of accountability can measure.
This review only scratches the surface of the myriad of micro forces that matter for productivity. What this says though is that the forces are complex, but they matter tremendously for macro productivity. The fact that these forces are complex suggests that there is no magic bullet, and no single policy reform by itself is likely to have a large effect.
References


Bravo-Biosca, A., C. Criscuolo and C. Menon (2013),“What Drives the Dynamics of Business Growth?”, OECD Science, Technology and Industry Policy Papers, No. 1,


Caselli, Francesco (2005), “Accounting for Cross-Country Income Differences,” Handbook of


Lagakos, David (2009), Superstores or Mom and Pops? Technology Adoption and Productivity Differences in Retail Trade,” working paper.

Levy, Santiago (2008), Good Intentions, Bad Outcomes, Brookings Institution Press.


