Firm Dynamics and Economic Performance

By

John Haltiwanger*

May 2011

*University of Maryland and NBER. This draft paper is for presentation at the ABCDE 2011 Conference in Paris.
I. Introduction

One of the most striking findings from the growing empirical literature on firm dynamics is that within narrowly defined industries there are large differences in productivity across businesses. This pattern holds in advanced and emerging economies. These differences reflect many factors such as managerial ability, organizational practices, choice of technology, vintage of the business and its capital, other intangible assets, and idiosyncratic productivity shocks. The evidence also shows that there are substantial differences across countries in terms of the within-industry dispersion of productivity as well as in the within-industry covariance between firm size and productivity. The working hypothesis is that these cross country differences reflect differences in the degree of misallocation across countries.

Misallocation can occur on many dimensions but in general refers to departures from allocative efficiency. Allocative efficiency refers, in turn to the extent to which resources are allocated to their highest valued use. In this context, there are both static and dynamic concepts of allocative efficiency. In terms of static allocative efficiency, more productive businesses should be larger. In terms of dynamic allocative efficiency, resources should be moving away from more productive to less productive businesses.

The greater availability of firm level data for many countries has shown that differences in economic performance (e.g., output per capita) across countries are closely related to differences in static and dynamic allocative efficiency across countries. The business and investment climate play a critical role in promoting or impeding such allocative efficiency.

Static allocative efficiency can obviously only be achieved if there is also dynamic allocative efficiency. Moreover, the challenge is that it takes time and resources to achieve both
static and dynamic allocative efficiency. The evidence suggests it requires a substantial rate of ongoing reallocation of output, capital and labor. Such reallocation is by its very nature costly to both firms and workers. Moreover, much can go wrong in these dynamics.

Part of the reallocation dynamics that is an inherent part of economic growth involves the entry and exit of businesses. Entering businesses face many challenges and obstacles. In addition, the evidence shows that entering businesses exhibit especially high dispersion of productivity. This reflects that new businesses are engaged in learning but also that some new businesses inherently discover they are ill suited to the activity with which they are engaged and depending on outside opportunities should potentially exit. Given the difficulties of starting up a business and the fragility of young businesses, distortions to the business climate can adversely impact the selection and learning dynamics of such businesses. In terms of the current discussion, an important margin over which distortions can have an adverse impact is on business start-ups and the post-entry dynamics of young businesses.

As noted, the reallocation of outputs and inputs across firms is costly – it is costly to businesses in terms of adjustment frictions and it is costly to households as workers are caught up in this reallocation (and also costly to households as they own the businesses incurring costs). Workers impacted by reallocation often spend time unemployed in transition and if at all prolonged often suffer substantial earnings losses. Substantial costs are born by businesses in terms of the time and resources used in accomplishing firm entry and exit as well as contraction and expansion. Some of these time and resource costs are an inherent component of the process of reallocation but market structure and institutions play a critical role in determining the extent to which the reallocation is productivity enhancing.
In this paper, I summarize the empirical and theoretical literature underlying the challenges of promoting allocative efficiency on the one hand and minimizing the disruption costs of ongoing reallocation. Section II provides an overview of the basic facts on firm dynamics. Section III presents conceptual underpinnings. Section IV discusses policy challenges. Section V provides some concluding remarks.

II. Basic Facts

A. Productivity and Reallocation

It is useful to start with basic facts about the distribution of productivity and size across businesses. Figure 1 illustrates a hypothetical productivity distribution within industries that reflects the patterns that have been observed in the data. As noted, a striking finding is that within narrowly defined industries there are large differences in productivity across businesses. For example, Syverson (2004) finds that the interquartile range within narrow U.S. manufacturing industries is over 30 log points for measures of establishment-level total factor productivity. Foster, Haltiwanger, and Syverson (2008) show that the dispersion of establishment-level total factor productivity within detailed product classes that abstracts from variation in plant-level prices is at least as large. Similarly, there is substantial dispersion in business size. Bartelsman, Haltiwanger and Scarpetta (2009a,b) show, for example, that firms in

---

1 In what follows, some of the evidence is about establishments and some of the evidence is about firms. By establishments, we mean specific physical locations where production activity is located. By firms, we mean all activity under common operational control. As an example, an individual Wal-Mart store is an establishment while the firm is the activity of all Wal-Mart stores as well as other establishments owned and controlled by Wal-Mart (e.g., distribution facilities). Both establishment and firm level evidence is relevant. For job reallocation, establishment-level is likely preferred since the frictions in the labor market are very much about moving workers away from one location to another. Note in addition that most establishment-level job reallocation is between firm reallocation. For other purposes, analyzing activity at the firm level is preferable. For example, in terms of discussing financial market frictions, the relevant level of activity is the firm not the establishment. The discussion attempts to be clear when the results are at the establishment level or at the firm level. Note that theoretical models often do not make this distinction – that is they don’t formally model multi-establishment firms.
the top quartile of the U.S. size distribution are on average 80 times larger than firms in the first quartile of the within industry size distribution.

The large dispersion of productivity and size provide ample scope for there to be differences across countries, time periods within countries and industries within countries in “static” allocative efficiency. By the latter we mean the extent to which in the cross section resources are allocated to their highest valued use which in this case implies that the most productive firms should be the largest firms. Figure 2 (based on Bartelsman, Haltiwanger and Scarpetta (2009a,b)) show there are large differences in the within industry covariance of size and productivity across countries. For example, the covariance in firm size and firm productivity in the U.S. is high and positive while it is lower in Western Europe and still lower in Eastern Europe. Interestingly, while the covariance between size and productivity is low in Eastern Europe it has been increasing substantially over the last couple of decades as seen in Figure 3. Bartelsman, Haltiwanger and Scarpetta (2009a,b) also show that these differences in the size/productivity covariances are potentially quite important in accounting for differences in output per capita across countries.

The covariance measures depicted in Figures 2 and 3 are a component of a productivity decomposition developed by Olley and Pakes (1996). The Olley-Pakes decomposition decomposes an index of industry level productivity into an unweighted mean of productivity at the firm level and the covariance of size and productivity. Figures 4 and 5 show Olley-Pakes decompositions of within industry productivity for Columbia (using TFP as the measure of productivity) based on study by Eslava et. al. (2004) and for China (using labor productivity) based on study by Deng and Haltiwanger (2008). Both countries underwent substantial market reform in the sample periods for these analyses. It is striking that in both countries the
covariance between size and productivity rose substantially. Moreover, it is especially striking that in 1998 the covariance between size and productivity in China was negative. The interpretation is that at that point the largest firms were relatively low productivity firms. Figure 5 suggests that an important part of China’s rapidly growing productivity is moving that covariance from negative to slightly above zero. A covariance at or around zero is still quite low relative to say the U.S. But this also suggests there remain considerable opportunities in China to improve allocative efficiency.

While the variation in the within industry cross sectional patterns of productivity and size across countries are of critical interest and importance, they offer an incomplete picture. That is, on the basis of the cross sectional evidence alone one might conclude that there is relatively stable within industry size and productivity distribution in the sense that high productivity firms remain high productivity firms and large firms remain large firms and so on. While there is persistence in both firm size and firm productivity, there also is considerable reallocation and movements within the distributions. Estimates of the persistence of idiosyncratic of productivity shocks suggest first order yearly autocorrelation of about 0.8 (see, e.g., Foster, Haltiwanger and Syverson (2008)). Along with estimates of dispersion, this estimate of persistence implies estimates of the standard deviation of innovations to productivity shocks of about 0.20 (in terms of log total factor productivity).  

Along with this high variance of innovations to productivity shocks, there is a high pace of reallocation of outputs and inputs. Figure 5 (based on Haltiwanger, Jarmin and Miranda (2010)) shows that the annual establishment-level gross job creation rate of about 18 percent (as a percentage of employment) and an annual establishment-level gross job destruction rate of 16 percent in the U.S. This implies in any given year a gross job reallocation rate of about 34

---

2 This statistic is consistent with the evidence in Foster, Haltiwanger and Syverson (2008).
percent --- that is about 34 percent of jobs are reallocated each year in the U.S. Figure 5 also shows that entry and exit of firms as well as entry and exit of establishments of existing firms play an important role in this reallocation. Bartelsman, Haltiwanger and Scarpetta (2009a,b) show that such patterns are present in a range of advanced and emerging economies. In addition, Davis and Haltiwanger (1999) and Haltiwanger, Jarmin and Miranda (2010) show that much of this reallocation is within industries (about 90 percent of job reallocation in the U.S. is within 6-digit NAICS or 4-digit SIC industries). Thus, it reflects the contribution of business entry, exit, expansion and contraction within industries.

Just as there is a relationship in the cross sectional distribution of size and productivity, there is a relationship between the pace of reallocation and productivity shocks. In well functioning economies, outputs and inputs are being reallocated away from the lower productivity to higher productivity businesses. Figure 6 shows that about a third of the productivity growth within a manufacturing industry over a 5 year period of time is accounted for by such reallocation in the U.S. (this is based on Foster, Haltiwanger and Syverson (2008)). Foster, Haltiwanger and Krizan (2001) show that over longer horizons (10 years) the contribution is even larger (about 50 percent).

An important component of this reallocation is entry and exit of establishments and firms. Given the importance of entry, it is instructive to characterize the post-entry dynamics of young firms. Figure 7 (based on Haltiwanger, Jarmin and Miranda (2010)) shows how job destruction and net employment growth at the firm level vary with age of the parent firm.3 Among surviving firms, young firms grow very fast in absolute terms and relative to their more mature counterparts. However, the employment-weighted exit rate (equivalent to the job destruction rate from firm exit) is also much higher for young firms. Taken together, the

---

3 These patterns show dynamics at the firm not establishment level.
implication is that young firms exhibit an “up or out” dynamic – they either grow fast on average or they exit.\(^4\)

How do these “up or out” dynamics relate to productivity? Figure 8 (based on Foster, Haltiwanger and Krizan (2006)) shows the relationship between productivity and continuing and exiting for all and single unit establishment firms in retail trade. Exiting young establishments and firms have very low productivity while surviving young firms and establishments have above average productivity. As such, the “up or out” dynamic in Figure 7 contributes to productivity growth through moving resources towards the more productive and away from the less productive young businesses.

In short, in well functioning economies there is evidence of not only static allocative efficiency (more productive businesses are larger) but dynamic allocative efficiency (resources are being moved from less to more productive businesses). A key theme in the remainder of the paper is that the extent to which a country exhibits patterns of both static and dynamic efficiency will depend on market structure and institutions.

\(B.\) \textit{The Impact on Workers}

As noted in the introduction, the ongoing reallocation is not costless with workers and businesses bearing substantial time and resource costs in accommodating the reallocation even if it productivity enhancing. Both types of resource costs need to be taken into account in evaluating the extent to which a country is achieving static and dynamic allocative efficiency.

\(^4\) A related key message of the Haltiwanger, Jarmin and Miranda (2010) paper is that firm age rather than firm size is the more theoretically and empirically relevant characteristic of businesses for job creation. That is, the conventional wisdom that small businesses create most jobs is better understood in terms of the job creation prowess of business startups and the rapidly growing surviving firms in Figure 7. Business startups and young firms are small so the conventional wisdom is picking up the role of startups and young firms. On the flip side, small, mature firms are not disproportionate creators of jobs.
In terms of the impact on workers, the evidence shows that in healthy times in healthy economies the impact on workers is not too adverse in terms of employment and earnings outcomes. For this purpose, we focus on the evidence in the U.S.\(^5\) Figures 9 and 10 (based on from Davis, Faberman and Haltiwanger (2011)) help highlight several key patterns. These figures show that in good economic times much but not all of job destruction in the U.S. is accommodated by worker quits instead of layoffs. This pattern is consistent with related evidence that shows that in good economic times, many separations of workers are associated with either no spell of joblessness or a short spell of joblessness and often result in an increase in earnings relative to the prior job. The latter is consistent with the perspective that the workers are reallocating away from lower productivity (and/or from the perspective of both the worker and the firm, a low quality match) firm to a higher productivity firm or match.

All of the potential problems with dislocation are significantly exacerbated in economic downturns even in otherwise healthy economies. Not surprisingly, in an economic downturn job destruction increases and job creation decreases. Job destruction in downturns is accommodated mostly through layoffs that yield spells of unemployment that are often protracted. The recent economic downturn in the U.S. offers ample evidence of these challenges. In normal times, the average duration of unemployment in the U.S. is about 2 months. In the recent economic downturn, it decreased to over 10 months. The evidence shows that the persistent earnings losses that workers who experience longer term unemployment are worse in recessions.\(^6\)

All of the above is for healthy, well-functioning economies. For highly distorted economies, reallocation is not well accommodated at any time. In highly distorted economies there is often an effort to stifle reallocation. One can understand why given the concerns about

---

\(^5\) See Davis, Faberman and Haltiwanger (2010) and references therein.
long term unemployment and the impact of displacement on earnings. However, as we discuss below, stifling such reallocation has adverse effects on static and dynamic allocative efficiency.

C. Churning for Churning’s Sake?

The discussion thus far has focused on the importance of flexibility given the role of productivity enhancing reallocation for economic growth and performance. But the discussion has also highlighted the costs of churning even for healthy, well functioning economies. Given the costs, churning for churning’s sake has no value – it is only if the churning reflects the push towards improved allocative efficiency that such churning has value. In this respect, it is not especially relevant or informative to simply compare the pace of churning across countries as an index of its dynamism and flexibility. Still it is natural to inquire about what we know about differences in the pace of creative destruction across countries.

Comparing the pace of churning or reallocation across countries has proven to be a substantial measurement challenge as well as conceptual challenge for reasons related to the discussion above. The working conjecture is that the U.S. being a very flexible economy would have a higher dispersion of growth rates of outputs and inputs (e.g., employment) than other countries. However, the evidence on this is mixed. Part of the reason for this is measurement difficulties (see, Bartelsman et. al. (2009a,b)). However, another reason might be flexibility manifesting itself in different dimensions. As Bertola and Rogerson (1997) emphasize, countries with rigid labor regulations also often have centralized wage bargaining. The former should dampen employment volatility while the latter should increase employment volatility.

A more promising approach towards understanding the impact of the business climate on the pace of reallocation has been to explore changes within countries in response to market reforms and/or to use some other variation within countries to identify the impact of the business
climate. For example, there is much evidence that the pace of reallocation increased dramatically in the 1990s in the transition economies (see, e.g., Faggio and Konings (1999), Jurajda and Terrell (2002) and Haltiwanger and Vodopivec (2003)). Another approach to identify the impact of the business climate (including policies promoting or deterring flexibility) is to use a difference-in-difference identification approach. For example, Haltiwanger, Scarpetta and Schweiger (2010) use differences in volatility across industry and size classes within countries to show that countries with more rigid labor markets have less employment reallocation.

III. Conceptual Underpinnings

D. Core Models of Firm Dynamics

We begin with canonical models of the determinants of the size distribution of activity, static allocative efficiency, dynamic allocative efficiency and firm and industry dynamics. One of the canonical models of the determination of firm size is based on assuming some form of decreasing returns is present given economies of scope and control (e.g., Lucas (1978)). Economies of scope imply that the most productive firm will not take over the market since as the most productive firm gets large decreasing returns set in. Another common model of the determination of firm size is to assume that firms face downward sloping demand curves – models of product differentiation such as those in Melitz (2003) (and many antecedents) have this feature. Such product variation need not be differences in physical products but can also include differences in the bundled goods and services of providing the good or service in question (including the location of providing the good or service). Product differentiation also implies the most productive firm will not take over the product since the goods produced b
different firms are not perfect substitutes. Whether due to decreasing returns or product differentiation, a core implication of these models is that more productive firms should be larger.

With such models as a backdrop, there are a rich set of models that help us understand the observed industry and firm dynamics. Jovanovic (1982) posits that at entry firms don’t fully know their productivity (or other aspects of profitability) and so an important part of firm dynamics especially for growing industries is the selection and learning dynamics of young firms. Those firms that learn they have a good location, good product or process, survive and grow. Those that learn they are not profitable contract and exit. Since the evidence on firm dynamics shows that reallocation and restructuring is not confined to young firms, additional theories need to be used to understand such dynamics. Ericson and Pakes (1996) (and a variety of others – see the recent survey by Syverson (2009)) develop models that help account for the ongoing reallocation and productivity dynamics. Ericson and Pakes (1996) postulate that every time a firm makes a major change in its way of doing business (either by adopting a new technology or in responding to some major change in economic conditions like higher energy costs), the firm gets a new draw on its profitability and productivity with associated selection and learning dynamics.

The more general notion as illustrated in models such as Hopenhayn (1992) and Hopenhayn and Rogerson (1992) is that the productivity shocks firms face are persistent but that firms are constantly subject to new productivity and profitability shocks. Viewed from this richer perspective, firms are constantly forced to adjust and adapt to changing economic circumstances and while their past successes can help in forecasting their ability to adjust and adapt, firms are constantly required to reinvent themselves. Those that reinvent themselves well survive and grow. Those that adapt and adjust poorly contract and exit.


E. Scope for Misallocation

Much of the above discussion paints a picture of the potentially important role of productivity enhancing reallocation for economic growth. More recent work has emphasized all of the many factors that can go wrong as countries try to achieve both static and dynamic allocative efficiency. Banerjee and Duflo (2003), Restuccia and Rogerson (2009), Hsieh and Klenow (2009) and Bartelsman, Haltiwanger and Scarpetta (2009a,b) all emphasize that there are a host of distortions to static and dynamic allocative efficiency. Such distortions include barriers to entry and exit, regulations that deter job destruction, poorly functioning product, capital and labor markets, weak rule of law, poor public infrastructure for communication and transportation, as well as problems with graft and corruption or otherwise arbitrary and capricious behaviour of governments. The consequences of such distortions can be severe. As discussed above, in an ideal setting the most productive firms are the largest firms. In a distorted economy with poor institutions, the largest firm may not be the most productive but rather the best connected or perhaps the best at navigating the distortions within a country.\(^7\)

This recent literature has shown that the misallocation that results from the type of distortions discussed can account for a substantial fraction of the observed differences in proxies for allocative efficiency (such as the size/productivity covariance discussed in section II) as well as accompanying differences in aggregate output and consumption per capita.

Such misallocation distortions have adverse consequences in their own right but also potentially yield a variety of second best problems for economic reforms. If it is difficult to start a business, difficult to expand, difficult to avoid having rents extracted from any profits unless one stays sufficiently small (or even informal), difficult to contract and/or exit (say due to poor

\(^7\) Bartelsman, Haltiwanger and Scarpetta (2009a,b) provide evidence on differences across countries on a wide range of distortions.
bankruptcy regulation and enforcement) and/or any number of other distortions, the productivity enhancing reallocation discussed above from market reforms can be derailed.

In like fashion, not only might the reallocation be derailed but it may be especially costly. As emphasized by Caballero and Hammour (2000), distortions can be such that creation and destruction get decoupled in time – that is, market reform might induce downsizing and exit by less productive businesses as appropriate but the accompanying creation and expansion by the more productive businesses may be delayed or derailed. When there is such decoupling, the cost to workers can be especially high since in an economy with lots of destruction but not much creation (at least for a period of time) there is by construction an economic downturn with many dislocated workers.

One caveat that has been expressed about the above arguments is that the role of reallocation for productivity growth may be more of an issue for advanced market economies than emerging economies. The argument that is made is that it is economies at the frontier of technology that are inherently engaged in the experimentation and creative destruction process. Following this reasoning, the argument for emerging economies is that if technology could simply be brought up to levels from the past in advanced economies where methods and business practices are well understood then this would be still be a substantial improvement. There are several reasons why this line of argument is not persuasive. For one, the evidence shows that in all economies (advanced and emerging) we observe large within sector differences in productivity across businesses (see, e.g., Bartelsman, Scarpetta and Haltiwanger (2009a,b), and Hsieh and Klenow (2009). If anything within sector dispersion in productivity is larger in emerging economies reflecting as Hsieh and Klenow (2009) emphasize the effects of misallocation. The point is there is much scope for productivity enhancing reallocation in
emerging economies. For another, while the sources of within industry differences in productivity across businesses are still under investigation, it is clear that they don’t simply stem from access to different “blueprints” for how to produce specific goods and services. Rather, differences in productivity reflect differences in managerial ability, organizational capital, management practices and other intangible factors (see, e.g., Corrado, Haltiwanger and Sichel (2005)) along with potentially random factors about choosing the right combination of location, products, and processes. The implication is that productivity differences across businesses reflect idiosyncratic factors that are not simply a matter of blueprints – and that such differences are pervasive in high tech and low tech sectors as well as advanced and emerging economies.

While this discussion highlights much progress has been made in our understanding of these issues theoretically and empirically, there are many open questions on these issues that are also active areas of research. Identifying the potential benefits in terms of improved allocative efficiency and the costs in terms of transition costs and worker dislocation from economic reforms is an active area of research.

**IV. Policy Lessons and Challenges**

The policy lessons in broad terms are clear but the actual implementation imposes many challenges. The broad policy lesson is that a healthy economy needs to be sufficiently flexible to permit productivity enhancing reallocation while minimizing the disruption costs from such reallocation in manner that does not stifle the reallocation. Few countries achieve the economic environment that is consistent with this broad lesson. One could argue the U.S. has the market structure and economic institutions that closely approximate this objective in healthy economic times. But the recent recession has reminded us that even in the U.S. there is fragility in the system and disruptions in key markets (like financial markets) disrupts the nature and
consequences of accommodating the economic volatility that is part of the ongoing process of making technological progress. So one of the policy challenges is how to maintain the market structure and economic institutions that operate in healthy economic times but then permit intervention when markets get disrupted.

For emerging economies, the challenges are potentially enormous. As Pages (2010) and Pages, Pierre and Scarpetta (2009) discuss in great detail, one great challenge evident in many emerging economies is the role of informality and in turn what they call the “missing middle”. In highly distorted economies where the burden of poor institutions and market structures weigh down on businesses, there tends to be very small businesses, very large businesses but not as many medium sized businesses as in healthy market economies. They argue the reason is that only the very large businesses have the resources to deal with the highly distorted economy (or worse are simply large because of the highly distorted economy – the businesses are well connected in some fashion). They argue that small businesses (even those with great potential in terms of productivity) stay small to essentially fly below the radar. That is, businesses stay small and informal so they are not regulated, taxed or as subject to graft and corruption since it is difficult to extract rents from such businesses.  

Reallocation has less chance of being productivity enhancing in such economic environments. Moreover, it is unclear that trade reform will have the effects discussed in prior sections in terms of either theoretical predictions or actual outcomes like those experienced in Colombia and Chile. Even in the latter countries, the evidence discussed above is about what happened to the formal establishments and firms in the manufacturing sector. It is certainly

---

8 There may an ameliorating effect on the duration of joblessness in economies with large informal sectors to the extent that workers dislocated by restructuring and reallocation can quickly find jobs in the informal sector. It is not clear that this is indeed beneficial to the extent it reflects workers and firms in the informal sector being underemployed for the reasons discussed in the text.
possible that the benefits discussed for formal firms (and the relatively modest adverse effects for formal sector workers) only apply to the formal sector. It would be quite interesting to explore how the informal sector fared in these countries over this same period.

The challenges then are that many components need to be in place for economies to successfully grow while opening up markets. The full list of components is long. Labor markets need to be sufficiently flexible to permit reallocating workers from less productive to more productive establishments without intervening long spells of unemployment. As part of this flexibility, safety nets need to be in place so that workers adversely impacted by reallocation can be assisted in finding new employment without distorting the process of reallocation. The infrastructure needs to be sufficiently of high quality to insure that existing and starting up businesses that seek to grow are not thwarted by factors such as poor transportation and communication. Product markets need to be sufficiently competitive that firms are not large for reasons of market power (or having obtained favourable treatment from the government). Financial markets need to be sufficiently developed to provide funding to starting up and expanding businesses and to be able to deal with the inevitable failure of young and small businesses. Regulation has to provide appropriate oversight without imposing onerous time and resource costs on starting up a business or shutting down a business. The legal system has to work sufficiently well so that property rights are well established and bankruptcy and business failure can be accommodated. The rule of law and the role of the government need to be such that graft, corruption and other forms of criminal activity don’t thwart private sector businesses from starting and growing (and becoming formal). These are just examples of the many components that need to be in place. With all of these components in place, opening up to
markets and competing in world markets is much more likely to be productivity enhancing without the costs of reallocation being too high for businesses and workers.

Getting all of these pieces in place simultaneously is obviously a challenge on many dimensions. Given such challenges, governments often try to intervene to facilitate growth and/or to protect workers and businesses from some of the adverse effects of the impact of volatility (some of which stem from opening up of markets to globalization). The message of this paper is that policies and institutions that stifle reallocation can yield very poor outcomes. Another related message of this chapter is that well-intended industrial policies that try to aid the private sector must confront the facts associated with the large dispersion of productivity across businesses (and the associated productivity enhancing reallocation that works in healthy market economies). Recall that dispersion of productivity in narrowly defined sectors in advanced economies like the U.S. is very large and even larger in less developed economies. Industrial policies that (perhaps inadvertently) support the low productivity businesses in a sector will lower aggregate productivity in a country and make it difficult for the country to increase its productivity over time (if, for example, it is difficult for governments to let go companies they have supported). The government is in no better position than the market to pick winners and given the evidence on dispersion, the risks of picking and supporting low productivity businesses is non-trivial. As an alternative to industrial policies, policies that seek to address the distortions and market failures in the country have much more promise.

Another challenge is how to handle crises. In crises, even in otherwise healthy economies reallocation dynamics get distorted. In crises there is lots of job destruction but not much job creation with accompanying high unemployment. In crises especially like the recent financial crisis, financial markets are not facilitating reallocating resources away from less
productive to more productive businesses.\textsuperscript{9} Such productivity enhancing reallocation requires in part at least financial markets providing funding to startups and to young, small businesses who have the potential to be high growth firms. This breaks down in recessions that are associated with financial crises.

V. \textit{Concluding Remarks}

A summary of the main messages of this paper is as follows:

- The evidence strongly supports the view that static and dynamic allocative efficiency as captured by the relationship between productivity and size in the cross section and productivity and resource reallocation over time are critical for aggregate economic performance of a country.

- Underlying this evidence are basic facts about the distribution of size and productivity on the one hand and ongoing resource reallocation and productivity on the other. In the cross section, we observe a very dispersed and skewed size distribution of activity in advanced market economies that is accompanied by a very dispersed and skewed distribution of productivity. In a well functioning economy, these two distributions should be strongly positively correlated – that is, the most productive businesses should be the largest businesses. In addition, in a well functioning economy, the reallocation of resources should be reallocating resources away from less productive businesses and towards more productive businesses. The evidence shows there is considerable variation across countries in the extent to which size and productivity are correlated and reallocation is productivity enhancing.

\textsuperscript{9} A recent paper that explores these issues is Eslava et. al. (2010c). They find that exits are less related to productivity in times of financial crises.
In a highly distorted economy on other dimensions, there are second best problems so that piecemeal policy reform will not be as effective in such distorted economies. Distortions may arise in the legal system and the rule of law as well as in regulation and in product, labor and financial markets. A poorly functioning labor market makes the response to reallocation very costly. Reallocation yields inherent costs on both businesses and workers as it induces workers to relocate across businesses. In a poorly functioning labor market, this can be very costly. Even in advanced market economies that are normally healthy, in severe economic downturns the reallocation dynamics of workers becomes distorted. Addressing how to combat the difficulties of managing reallocation dynamics during economic downturns without distorting the potential for productivity enhancing reallocation in the long run is a continuing challenge.

Well functioning financial markets play a critical role in facilitating static and dynamic allocative efficiency. A feature of healthy advanced market economies is they are constantly reinventing themselves as businesses and households adapt and adjust to changing economic conditions and market opportunities. Part of this reinvention process involves new firms entering and exploring new products, processes and ways of doing business. For each entering cohort, many businesses fail in the first five to ten years. However, conditional on survival, young businesses grow faster than their more mature counterparts. In addition, among the young businesses are high growth businesses that contribute disproportionately to innovation, job growth and productivity. Financial markets need to be sufficiently well developed and functioning to help provide the financing to startups and high
growth young businesses as well as being capable of absorbing the exit of low productivity businesses.

- The recent financial crisis highlights how this process can break down and distort reallocation dynamics. In times of financial crises, financial markets are less able to facilitate the selection and growth dynamics of businesses – for large, mature as well as young and small businesses alike. Perhaps ironically the globalization of financial markets has made the problem more challenging during economic crises given the flights to quality that increasingly spreads globally during crises. Financial regulation that helps monitor the health of the financial services industry and provides safeguards against financial collapses are undoubtedly needed. Some caution about how to design such safeguards is provided by the underlying message of this chapter.

The successful new, young firms need equity investors and the development of venture capital, angel financing, and other such markets that target startups and young and small businesses has facilitated productivity enhancing reallocation. The message then is that financial sector reform should avoid increasing the barriers to the financial sector in finding new instruments and creative ways of providing funding to high growth businesses and more generally to productivity enhancing reallocation.

- The recent economic crisis has also highlighted the potential importance of heightened uncertainty during economic crises being a significant damper on economic recovery from such crises. The key insight from economic theory that has empirical support especially in the recent crisis is that heightened uncertainty will slow down recoveries due to caution and waiting effects. That is, even businesses with profit making opportunities will delay and/or reduce the amount of investment
and hiring due to heightened uncertainty. Such adverse effects of uncertainty are clearly relevant for all economies as the recent crisis has shown but are likely especially important in emerging economies that inherently have a higher degree of uncertainty at all times. One of the challenges of economic reform including trade reform is to address the impact of heightened uncertainty due to economic crises as well as due to the market reforms themselves.
References:


Preliminary and Incomplete – Do Not Cite Without Permission


**Figure 1**

The Distribution of Productivity Across Businesses in the Same Industry

Interquartile range WITHIN narrow industries is over 30 log points

Notes: The above is a hypothetical depiction of the shape of the productivity distribution reflecting the empirical finding that productivity is approximately log normally distributed. The reported interquartile range is based on the distribution of U.S. manufacturing establishments (see, Syverson (2004)).
Figure 2: The Relationship Between Size and Productivity

A. The Average Within-Industry Covariance Between Size and Productivity, 1992-2001

![Covariance Between Size and Productivity](chart1)


![Change in Covariance](chart2)

Figure 3: Olley Pakes Decompositions of Productivity for Colombia

Source: Eslava et. al. (2008). This is a decomposition of within industry TFP growth.
Figure 4: Olley-Pakes Decomposition for China

Olley-Pakes Decomposition of Labor Productivity
(Average Industry)

Source: Deng and Haltiwanger (2008)
Figure 5

**Job Creation and Destruction, U.S. Private Sector, Annual Rates (Percent of Employment), 1980-2009**

- **Job Creation**
  - Continuing Establishments: 10.3
  - New Establishments (Existing Firms): 3.2
  - New Firms: 3.1

- **Job Destruction**
  - Continuing Establishments (Continuing Firms): 10
  - Exiting Establishments (Continuing Firms): 2.4
  - Exiting Firms: 2.9

Source: Tabulations from the BDS combined with tabulations from Haltiwanger, Jarmin and Miranda (2010).
Figure 6

Components of Total Factor Productivity Growth over Five-Year Horizons, 1977-1997, Selected Manufacturing Industries

Source: Foster, Haltiwanger and Syverson (2008)
Figure 7

Up or Out Dynamics of Young U.S. Firms

Source: Haltiwanger, Jarmin and Miranda (2010)
Figure 8

**Productivity of Young Businesses Relative to Mature Surviving Incumbents, U.S. Retail Trade**

<table>
<thead>
<tr>
<th></th>
<th>Single Unit Establishment Firms</th>
<th>All establishments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Exits</td>
<td>-31.60%</td>
<td>-32%</td>
</tr>
<tr>
<td>Mature Exits</td>
<td>-26.20%</td>
<td>-27%</td>
</tr>
<tr>
<td>Young Survivors</td>
<td>2.80%</td>
<td>3%</td>
</tr>
<tr>
<td>Young Survivors Five Years Later</td>
<td>1.20%</td>
<td>5%</td>
</tr>
</tbody>
</table>

**Source:** Tabulations from Census of Retail Trade taken from Foster, Haltiwanger and Krizan (Table 5, 2006);

**Note:** Young establishments are those that are under five years old.
**Figure 9. Quits, Layoffs, and Job Destruction in U.S. Private Sector**

Source: Figure 1 from Davis, Faberman and Haltiwanger (2011).

**Figure 10. Hiring and Job Creation in U.S. Private Sector**

Source: Figure 2 of Davis, Faberman and Haltiwanger (2011).