4 Globalization and economic volatility

John Haltiwanger*

4.1 Introduction

Businesses and households face substantial idiosyncratic and aggregate economic risk. As a general principle, economic risk for businesses reflects the myriad of factors that impact the profitability of the business, while for individuals economic risk reflects the myriad of factors that impact the earnings and employment outcomes of household members. While aggregate risk gets most of the headlines, the volatility of profitability and income that an individual business or household faces is dominated by idiosyncratic risk. That is, of the plethora of economic shocks impacting the outcomes for households and businesses, the evidence shows that the variance of idiosyncratic shocks is at least an order of magnitude larger than the variance of aggregate shocks.1 For example, whether a business is profitable reflects primarily idiosyncratic factors such as product quality, product mix and choice of technology, broadly defined, including the choice of business organization, factor mix, location and business-specific productivity, and cost and demand factors. Similarly, for households, earnings and employment outcomes primarily reflect the education and skills of household members as well as whether household members are well matched in the labour market.

Not only is idiosyncratic risk of critical importance at the micro level, but also recent evidence has highlighted that the manner in which an economy manages the idiosyncratic risk that households and businesses face plays a critical role in aggregate outcomes. That is, aggregate income and productivity in a country depends critically on how well the economy manages idiosyncratic risk.

In this chapter, we focus on idiosyncratic risk and associated volatility. The nature of how economies manage idiosyncratic risk is closely linked to how well they manage changes in economic conditions. Globalization is one of the core factors behind changing economic conditions and the impact of globalization on a country is closely linked to how well it manages idiosyncratic risk.

* I thank Marc Bacchetta, Marion Jansen and anonymous referees for comments on an earlier draft of this chapter.
What underlies this idiosyncratic risk to households and businesses? There are many factors, but the evidence shows that one key factor is that there are large differences in productivity across businesses even within narrowly defined sectors. Moreover, while these differences in productivity are persistent, there is a process of continuous change in the distribution of productivity. In addition to this dispersion in productivity, in advanced market economies there is a high pace of ongoing reallocation of outputs and inputs across businesses. In healthy market economies, the nature and pace of reallocation is closely tied to the distribution of productivity, that is, outputs and inputs are being reallocated away from less-productive businesses to more-productive businesses. In that respect, reallocation and the accompanying volatility at the firm level has the potential to enhance both productivity and welfare. In healthy market economies (and in healthy times in such economies) the evidence shows that a large fraction of aggregate productivity growth is associated with this ongoing reallocation. One needs to be careful about making causal inferences here – it is not reallocation per se that yields productivity growth but rather the process of productivity growth requires ongoing productivity-enhancing reallocation. The reason is that there is need for experimentation and trial and error in developing new products and processes, as well as in adapting to changes in the economic environment.

However, by its very nature 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 because workers are caught up in this reallocation and also because households own the businesses incurring costs. Workers impacted by reallocation often spend time unemployed in transition, and if this unemployment is at all prolonged they often suffer substantial earnings losses. Substantial costs are borne 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 enhances productivity.

How does globalization fit in with these dynamics? Globalization is one of the core factors that induce reallocation – in principle, the opening up of markets and the reduction of trade barriers permits productivity-enhancing restructuring and reallocation. The traditional view is that this permitted increased specialization into the production of products for which a country has comparative advantage. While there is some truth to this traditional wisdom, the development of rich new firm-level data that tracks trade flows at the firm level across countries (as well as at the detailed product level across countries) highlights the enormous amount of within-sector trade flows between countries. We have learned that exporting is rare at the firm level and the distribution of trading activity among firms that do export is highly
skewed. That is, most exporting firms export only a small number of products to a small number of countries. However, the bulk of trade is accounted for by the larger firms that trade in many products to many countries. In addition, we have learned that it is the most-productive firms that are engaged in trade.\(^3\)

These firm-level trade patterns are potentially consistent with productivity-enhancing reallocation but there are potential pitfalls and caveats. First, as an economy becomes more open the transition period can involve substantial dislocation of businesses and workers with the associated costs discussed above. Second, both during the transition and as a feature of longer-run outcomes, poor market structure and institutions can act as a barrier for productivity-enhancing reallocation. Put differently, trade liberalization in an economy with many market distortions can yield especially adverse outcomes and perhaps few benefits.

Globalization also involves globalized financial markets. The increased sophistication and globalization of financial markets is again in principle favourable for productivity-enhancing reallocation. That is, amongst other things, the ongoing need for reallocating outputs and inputs from less-productive to more-productive businesses involves firm entry, firm exit, firm expansion and firm contraction. Financial markets need to be working well to allocate credit to the business start-ups and expanding businesses. Since start-ups and young businesses tend to be more experimental, thus causing them to be more volatile, the financial markets must be able to manage and accommodate not only the start-ups and expanding businesses but also the high probability of contraction and business exit. Globalization has contributed to the development of richer markets with public trading of equity funds across the globe as well as the development of hedge funds, venture capital funds and private equity funds that not only operate in advanced economies but also in emerging markets. Such richer financial markets in principle yield better allocation of financial risk through diversification and the richer financial instruments available. However, it is also clear, especially from the past few years, that global financial markets are fragile and subject to sudden collapses in some segments which can become contagious in other segments of the market. Such fragility in financial markets can act as a source of undesirable volatility and a distortion to productivity-enhancing reallocation. Put simply, when financial markets break down, a business may contract or shut down not so much because it is a low-productivity business but because financial markets are no longer able to allocate credit to even potentially profitable businesses.

In this chapter, we summarize the theoretical and empirical literature underlying the challenges of promoting allocative efficiency, on the one hand, and minimizing the disruption costs of ongoing reallocation, on the other. Following from this we discuss the role of globalization in this context. Finally, we discuss the policy challenges of addressing the issues related to globalization and economic volatility.
The chapter is structured as follows: section 4.2 provides an overview of the basic facts on firm dynamics; section 4.3 presents conceptual underpinnings; section 4.4 gives a synopsis of the empirical evidence relating globalization to economic volatility and section 4.5 discusses policy challenges. Section 4.6 provides some concluding remarks.

4.2 Basic facts

Productivity and reallocation

It is useful to start with basic facts about the distribution of productivity and size across businesses. There is much evidence that even within narrowly defined sectors there is substantial dispersion in both productivity and size of businesses. For example, Syverson (2004) shows that the interquartile range of measures of within-industry establishment-level total factor productivity is about 30 log points. Foster et al. (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 et al. (2009a,b) show, for example, that in the United States within-industry firms in the top quartile of the 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 industry, countries, and 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. Bartelsman et al. (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 United States 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. Bartelsman et al. (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.

While the variations 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 productivity shocks suggest first order yearly autocorrelation of about 0.8 (see, for example, Foster et al., 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. Haltiwanger et al. (2010b) estimate an annual establishment-level gross job creation rate of about 18 per cent (as a percentage of employment) and an annual establishment-level gross job destruction rate of 16 per cent in the United States. This implies in any given year a gross job reallocation rate of about 34 per cent – that is about 34 per cent of jobs are reallocated each year in the United States. They also show that most of the establishment-level job reallocation is between firms and not between establishments within firms. Bartelsman et al. (2009a,b) show that such patterns are present in a range of advanced and emerging economies. In addition, Davis and Haltiwanger (1999) and Haltiwanger et al. (2010b) show that much of this reallocation is within industries (about 90 per cent of job reallocation in the United States is within 6-digit NAICS (North American Industry Classification System) or 4-digit SIC (Standard Industrial Classification) 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. The evidence suggests that about half of the productivity growth within a manufacturing industry over a ten-year period of time is accounted for by such reallocation in the United States (see Foster et al., 2001). In sectors like the retail trade, the evidence shows an even larger fraction of productivity growth is accounted for by reallocation (ibid.). The extent to which reallocation enhances productivity also varies across countries (see Bartelsman et al., 2009a,b).

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 chapter is that the extent to which a country exhibits patterns of both static and dynamic efficiency will depend on market structure and institutions. Moreover, for current purposes we are especially interested in how globalization
impacts the relationship between productivity and size in the cross-section as well as
the relationship between productivity and reallocation.

In the remainder of this chapter, we focus on economic volatility within industries. We do this not only because within-sector reallocation is much larger than between-sector reallocation but also because the literature has not found much impact of globalization on between-sector reallocation (see Goldberg and Pavcník, 2007). This latter finding is a bit of a puzzle which deserves further investigation.

One theme emphasized in this chapter is that accommodating micro volatility as evidenced by the ongoing need to reallocate workers to more efficient producers becomes disrupted in economic slumps. The nature of this disruption will be elaborated on in future sections. In addition, micro volatility can change the nature of macro volatility. For example, periods of intense restructuring in the economy can dampen aggregate activity as resources are being used for restructuring and reallocation rather than current production. In a related fashion, periods of intense restructuring are often associated with periods of heightened uncertainty which can slow down the adjustment dynamics from both aggregate and micro shocks. These relationships are also discussed in subsequent sections.

The impact on workers

As noted in the introduction, ongoing reallocation is costly, with workers and businesses bearing substantial time and resource costs in accommodating the reallocation even if it does enhance productivity. 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.

In terms of the impact on workers, the evidence shows that in healthy times in healthy economies the impact of reallocation on workers is not too adverse in terms of employment and earnings outcomes. For this purpose, we focus on the evidence in the United States.⁹ In good economic times in the United States, many reallocations of workers are associated with either no period of unemployment or a short period of unemployment 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 a lower-productivity firm (and/or from the perspective of both the worker and the firm, a low-quality skills match) to a higher-productivity firm or higher-quality match.

Also consistent with these patterns is that much but not all of the job destruction in the United States is accounted for by worker quits instead of lay-offs in good economic times, although in such times there are always some firm shutdowns with
accompanying worker lay-offs. Moreover, workers who experience a lay-off often have at least a spell of unemployment, and workers who separate from distressed firms via lay-offs and unemployment often have persistent earnings losses.

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 achieved mostly through lay-offs that yield spells of unemployment that are often protracted. The current economic downturn in the United States offers ample evidence of these challenges. In normal times, the average duration of unemployment in the United States is about two months. In the current economic downturn, it is closer to 10 months. The evidence shows that the persistent earnings losses for workers who experience longer-term unemployment are worse in recessions.10

All of the above conditions apply to 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 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.

4.3 Conceptual underpinnings

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 (for example, Lucas, 1978). 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). That is, it can be horizontal product differentiation rather than vertical product differentiation.

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 do not 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 that they have a good location, good product or process, survive and grow. Those that learn that 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 (1995) (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 (1995) 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 (1993) 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, they are constantly required to reinvent themselves. Those that reinvent themselves well, survive and grow. Those that adapt and adjust poorly, contract and exit.

Globalization potentially plays a key role in these dynamics. As Melitz (2003) and subsequent models emphasize, trade liberalization will induce a shake-up in the allocation of activity within an industry within a country. Melitz (2003) emphasizes that trade will permit the most productive and profitable firms to further expand which in turn will drive up factor prices (or potentially drive down mark-ups as in Melitz and Ottaviano, 2008) so that marginal firms in the industry will exit. The insight from this literature is that globalization can contribute to improved productivity within industries within countries as it induces productivity-enhancing reallocation. Of course, even productivity-enhancing reallocation it is not without costs, for all the reasons discussed above.

**Scope for misallocation**

Much of the above discussion paints a picture of the potentially important role of productivity-enhancing reallocation for economic growth and even how globalization can contribute to such 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 (2005), Restuccia and Rogerson (2008), Bartelsman et al. (2009a,b) and Hsieh and Klenow (2009) 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 labour markets; weak rules of law; poor public infrastructure for communication and transportation; as well as problems with graft and corruption or the 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.11

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 4.2) 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 including the potential benefits from trade liberalization. While the model of Melitz (2003) and related models make a case as to why liberalization can yield productivity-enhancing reallocation, in the presence of these distortions the impact of piecemeal economic reforms is less clear. 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 bankruptcy regulation and enforcement) and/or any number of other distortions, the productivity-enhancing reallocation highlighted by Melitz (2003) and others 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 (including trade 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. First, the evidence shows that in all economies (advanced and emerging) we observe large within-sector differences in productivity across businesses (see, for example, Bartelsman et al., 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. Second, while the sources of within-industry differences in productivity across businesses are still under investigation, it is clear that they do not 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, for example, Corrado et al., 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 that 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.

**Different dimensions of volatility**

Much of the discussion about volatility has focused on two dimensions of volatility. First, there is the large dispersion of productivity/profitability across businesses. Second, there is the ongoing reallocation of outputs and inputs across businesses. In terms of the latter, it is useful to note that such reallocation reflects an important form of dispersion across businesses – specifically, dispersion in output and input growth rates across businesses. That is, reallocation reflects resources from contracting businesses (those with negative growth rates in outputs and inputs) being reallocated to expanding businesses (those with positive growth rates in outputs and inputs). Entry and exit rates are at the extremes of the output and input growth rate distributions and obviously by construction contribute substantially to volatility.

It is natural to focus on dispersion in profitability/productivity on the one hand, and dispersion in output and input growth rates on the other hand. The core models
discussed in the first subsection above (“Core models of firm dynamics”) largely treat the dispersion in productivity/profitability as exogenous while treating the dispersion of output and input growth rates as endogenous. As highlighted in the discussion of the first and second subsections above, a critical factor impacting aggregate outcomes is how well an economy accommodates the idiosyncratic productivity/profitability shocks – that is, are those with favourable shocks growing and those with less favourable shocks shrinking, and in turn is such reallocation accomplished without too much disruption?

There are other closely related dimensions of volatility. An obvious closely related dimension is dispersion in earnings across workers. It is well known that in advanced economies there has been an increase in the dispersion of the level of earnings across workers – and the evidence suggests this is associated with changing technology favouring more-skilled workers (that is, skill-biased technological change) as well as closely related changes in trade patterns (the offshoring of lower-skilled jobs). This rise in earnings inequality is closely related to the firm dynamics discussed in prior sections. For example, a number of studies (for example, Davis and Haltiwanger, 1991; Dunne et al., 2004; and Barth et al. 2010) have found that much of the increase in earnings inequality in the United States is associated with an increase in the between-establishment dispersion in earnings. Moreover, these studies show that the establishments with the higher earnings are the more productive, more highly skilled and more likely to have adopted advanced technology.

What do we know about changes in volatility over time as well as difference in volatility across countries? Differences across countries as well as differences within countries over time in these different dimensions of volatility may reflect many factors. Differences may reflect changes in the driving forces (such as the factors driving dispersion in productivity/profitability) as well as changes in the adjustment dynamics. For the latter, an important issue in the current context is whether the differences reflect the relative flexibility of an economy and over what dimension. Greater flexibility might take many different forms. It might be that workers in a more flexible economy are more geographically mobile so that there is even more reallocation of labour in response to a given set of shocks. Alternatively, it might be that wages become more flexible (for instance with greater reliance on flexible pay mechanisms) so that a given set of shocks is reflected more in wages than in the reallocation of employment. The implication is that appropriate caution is needed in assessing differences in measures of volatility across time and across space.

In terms of the evidence of changes in volatility over time, the evidence is primarily for the United States which has longitudinal panels of businesses and workers over many decades to assess these issues. For the United States, there is evidence that volatility of output and employment growth rates of publicly traded firms has
increased for many decades (see, for example, Comin and Phillippon, 2006). However, interestingly, when the entire economy is considered (in the United States, publicly traded firms account for about 30 per cent of employment and 40 per cent of output), there is actually a pronounced decline in the volatility of employment growth rates (see Davis et al., 2007 and 2010a). Does this imply that the United States has become less flexible over time? This is an open research question but there is some evidence that it may reflect a different form of flexibility. Lemieux et al. (2009) show there has been a pronounced increase in the use of flexible pay mechanisms (bonus pay, stock options, and so on) in the United States, so this may reflect increased earnings flexibility. However, the evidence in Davis et al. (2007) suggests this is unlikely to be the whole story. For example, they find that this in part reflects the increasing shift in sectors like retail trade to large, national firms (for example, Wal-Mart) that are much less volatile than small family retailers. There is evidence that the shift to large, national chains reflects the type of technological change and reallocation discussed in the previous sections as large, national chains have been able to take greater advantage of advances in information technology for distribution networks and inventory control. However, it may also be that large, national chains are less nimble in adjusting to changing economic conditions. The more general point is that a decline in the pace of volatility in the United States may reflect a less dynamic US economy (which is thus less able to respond to changing economic conditions).

In terms of changes in the pace of volatility in other countries, there is much evidence that the pace of volatility increased dramatically in the 1990s in the transition economies (see, for example, Faggio and Konings, 1999; Jurajda and Terrell, 2002; and Haltiwanger and Vodopivec, 2003). It was clear this was disruptive with adverse aggregate consequences as most transition economies experienced a downturn in aggregate economic activity. Moreover, the evidence suggests that there was a non-trivial lag between the burst of job destruction and job separations early in the reforms and the subsequent recovery of job creation and hires. The patterns exhibited in the transition economy were consistent with the discussion and concerns about decoupling of job creation and destruction in the second subsection above (“Scope for misallocation”). Still, the evidence is that for the most part the transition economies weathered this storm and recovered with robust growth. It probably helped that the world economy exhibited robust growth in the second half of the 1990s.

Another issue of importance in terms of changes in the pace of volatility over time within countries is that periods of more intense restructuring are often associated with periods of heightened uncertainty. Bloom (2009) has stressed that this is important for understanding why business cycle downturns and recoveries differ due
to differences in the extent of uncertainty, Bloom et al. (2010) have emphasized that the Great Recession of 2007–09 is a period of especially heightened uncertainty due to the collapse of financial markets and the accompanying intense period of restructuring associated with this downturn (for instance, shifts away from construction activity and the restructuring of financial markets). Such heightened uncertainty contributes to especially slow recoveries since even businesses with potential profit-making opportunities are reluctant to invest and hire new workers due to the lingering heightened uncertainty during such crises.

In terms of evidence on differences in the pace of volatility across countries, this has proved to be a substantial measurement challenge as well as conceptual challenge for reasons related to the discussion above. The working conjecture is that the United States, being a very flexible economy, would have a higher dispersion of growth rates of outputs and inputs (for example, 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 labour regulations also often have centralized wage bargaining. The former should dampen employment volatility while the latter should increase employment volatility.

This discussion of different dimensions of volatility highlights the difficulties of simply comparing measures of volatility across countries or across time. As discussed in earlier parts of this chapter and in the next section, one approach that overcomes the measurement and conceptual challenges of comparing measures of volatility is to focus on whether the volatility (reallocation) enhances productivity. Differences across time and across countries on whether reallocation is productivity enhancing is of unambiguous importance. This is not to imply that measuring and studying differences in volatility across countries and time is not of interest or importance, but rather that the many different factors discussed in this section need to be taken into account. Another approach to identifying the impact of the business climate (including policies promoting or deterring flexibility) is to use a difference-in-difference identification approach. For example, Haltiwanger et al. (2010b) use differences in volatility across industry and size classes within countries to show that countries with more rigid labour markets have less employment reallocation. One can identify this effect not with the cross-country variation but with the within-country variation between industries and size classes.
4.4 What is the evidence on the impact of trade liberalization on productivity-enhancing reallocation and earnings and employment?

**Productivity-enhancing reallocation**

The discussion thus far has been broad-based in terms of the factors impacting productivity-enhancing reallocation and the potential adverse impact of reallocation on workers. That discussion helps provide the perspective to consider the direct evidence on the impact of trade liberalization on productivity.

Our focus is on the impact of trade on productivity-enhancing reallocation. However, before turning to that issue, it is useful to note that there is a large related literature that explores the impact of trade reform on the productivity of incumbent producers. A number of papers find that the productivity of incumbents increases after trade opening, including: Levinsohn (1993) for Turkey, Harrison (1994) for Côte d’Ivoire, Tybout and Westbrook (1995) for Mexico, Pavcnik (2002) for Chile, Trefler (2004) for Canada, Topalova (2004) for India and Fernandes (2007) for Colombia. We note, however, that De Loecker (2007) corrects for unobserved prices and finds that the impact of trade on productivity halves when controlling for prices and mark-ups rather than using standard productivity measures as the above studies. Also, Lileeva and Trefler (2010) have recently shown that gains in labour productivity from trade opening in Canada were concentrated in low-productivity firms that were induced by tariff cuts to start exporting. In these studies, the precise mechanism of how trade improves within-plant and within-firm productivity is typically not identified. It might be that opening to trade enables access to richer technologies (broadly defined) and/or opening to trade increases competitive pressures.

For our purposes, we are especially interested in papers that link trade reform, reallocation (volatility) and productivity. As noted, Pavcnik (2002) has a seminal paper on this topic using high-quality establishment-level data for Chile. Pavcnik is able to track longitudinal establishment dynamics of outputs, inputs and productivity following trade reform in Chile. She finds evidence that trade reform improves within-plant productivity and also evidence that trade reform improves allocative efficiency. She also finds that trade reform in Chile is associated with increases in the size/productivity covariance that contributes substantially to productivity.

Recent work by Eslava et al. (2010b) elaborates further on the insights from Pavcnik. Using high-quality longitudinal establishment-level data for Colombia, this work explores a number of channels through which trade liberalization impacts productivity. A core feature of this work is that the measures of total factor productivity abstract from the confounding of productivity and price effects that are a feature of much of
this literature. While this may seem to be a technical detail, it is important since it may be that trade reform impacts mark-ups and as such what looks like an increase in productivity might actually be an increase in establishment-level mark-ups.

Eslava et al. (2010b) find that trade reform in Colombia increased productivity through several channels. They find that trade reform increased the likelihood that low-productivity establishments exit – a pattern consistent with the predictions of the recent models on misallocation distortions. This improved market selection contributes positively to aggregate productivity. They also find the size/productivity covariance improves with trade reforms and the within-establishment productivity growth increases.

Between these two studies using high quality longitudinal establishment data for Chile and Colombia respectively, there is evidence in favour of the hypotheses that trade liberalization can improve productivity through improved allocative efficiency. However, some caution needs to be applied given that Chile and Colombia also engaged in other market reforms that accompanied trade reform. While these studies control for these other reforms, it may be that the other reforms permitted trade reform to work. Put differently, it may have been that the second-best problems discussed above were ameliorated in these two countries. In addition, these studies do not address the costs of reallocation including the impact on workers. We turn to this topic in the next subsection.

There is also direct evidence on the relationship between trade reform and volatility. Haltiwanger et al. (2004) present evidence that the trade reforms in Latin America in the 1990s systematically increased the pace of job reallocation in Latin America over this period of time. This evidence is consistent with the more detailed within-country studies discussed above but applies to a wider range of countries.

**Worker earnings, employment and dislocation**

There is a large literature looking at the distributional effects of trade liberalization (see Goldberg and Pavcnik (2007) for a complete review of this literature). The main focus of this literature has been on the skilled/unskilled wage differential since standard trade theory suggests that unskilled wages should increase in countries abundant in unskilled labour. Contrary to this, however, most studies find an increase in the skill premium in developing countries (for example, Borjas and Ramey, 1995; Robbins, 1996; Attanasio et al., 2004). In addition, a number of studies have found that trade liberalization is associated with a decline in wage premiums and an increase in income volatility (for example, Borjas and Ramey, 1995; Revenga, 1997; Goldberg and Pavcnik, 2004; Krebs et al., 2005). Few studies focus on the impact of trade liberalization on unemployment and households. The study by Attanasio et al.
(2004) is the only study to examine the relation between trade barriers and the likelihood of unemployment and they find no evidence of any relation. However, other studies have focused on the impact of trade protections on employment and the quality of employment. Most studies find a reduction in employment and, in particular, formal employment in sectors affected by trade liberalization (Borjas and Ramey, 1995; Revenga, 1997; Goldberg and Pavcnik, 2003). In considering these studies, note that their focus is on distributional effects (that is, which sectors or types of workers may be adversely impacted by trade) rather than on long-run economy-wide effects.

What do we know about what happened in terms of worker dislocation in these countries? Eslava et al. (2010a) explore what happened to workers in Colombia over the same period that they explored what happened to firm and aggregate productivity. They find evidence that, in general, worker dislocation has adverse effects on earnings and employment for workers who find themselves separated from a bankrupt firm. In that respect, the positive findings Eslava et al. find on improved market selection need to be balanced with the difficulties that workers face in separating from a bankrupt firm. Eslava et al. also find that the adverse effects of dislocation are not that persistent, and find only modest evidence that it is the workers in sectors with the greatest trade reform that are adversely affected. They also find evidence that is consistent with the literature on employment and earnings discussed above – that is, they find evidence that workers in sectors impacted by trade reform have lower earnings and formal sector employment and that these effects are larger for low-skilled workers.

One area of inquiry that would be useful to explore is the impact of globalization on the volatility that occurs during economic crises and in turn how workers impacted by volatility fare in globalized markets. As highlighted above in section 4.2, even in the United States job destruction has much more adverse impact on workers in economic downturns. We do not have extensive evidence on what happens to workers in economic downturns in emerging economies. However, some of the insights from the existing literature discussed in the previous sections sheds light on these questions. The evidence for transition economies suggests that wide-ranging and rapid market reforms yield an increase in the pace of restructuring and reallocation that can be quite disruptive. Virtually all of the transition economies suffered an economic downturn during the period of economic reforms. Those with rapid reform experienced sharp rises in job destruction and unemployment. However, after a period of adverse effects, the rapid reformers recovered, and an important part of that recovery was that the higher pace of restructuring was achieved through job-to-job flows (with lower rates of unemployment). At the time there was much debate about whether rapid or gradual reform made more sense – both theory and evidence provided support for both sides of this debate.
More recently much attention has been given to the role of heightened uncertainty during downturns associated with both financial collapse and intense restructuring. The theoretical and empirical evidence on these issues seems of particular importance for emerging economies. The key insight from this work is that in the recovery from downturns during periods of heightened uncertainty, businesses with profit-making opportunities are less likely to invest in capital and hire new workers. Such effects are always likely to be more important in emerging economies given the inherently greater uncertainty about the business climate in emerging economies. Also, in times of economic crises (perhaps especially those associated with a collapse in credit markets) the cautionary and delaying effects of uncertainty are likely to be that much more relevant in emerging economies. Exploring these hypotheses for emerging economies in greater detail should be a high priority for future research.

4.5 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 United States has the market structure and economic institutions that closely approximate this objective in healthy economic times. However, the recent great recession has reminded us that even in the United States there is fragility in the system, and disruptions in key markets (like financial markets) disrupt 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. This challenge of countercyclical policy is not the primary focus of this chapter but we discuss some issues along these lines below.

For emerging economies, the challenges are potentially enormous. As Pagés (2010) and Pagés et al. (2009) discuss in great detail, one great challenge evident in many emerging economies is the role of the informal sector and what they call the “missing middle”. In highly distorted economies where the burden of poor institutions and market structures weigh down on businesses, there tend to be very small businesses, very large businesses but not as many medium-sized businesses as in healthy market economies. Pagés et al. argue that 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.\textsuperscript{12}

Reallocation has little chance of enhancing productivity 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 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. Labour 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 of sufficiently 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 do not 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 enhance productivity 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 levels. 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 the opening up of markets to globalization). The message of this chapter 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 United States 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 of companies they have supported). The government is in a no better position than the market to pick winners and, given the evidence on dispersion, the risks of picking and supporting low-productivity businesses are not 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 a lot 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. Such productivity-enhancing reallocation requires, at least in part, financial markets to provide funding to start-ups and to young, small businesses that have the potential to be high-growth firms. This breaks down in recessions that are associated with financial crises.

### 4.6 Concluding remarks

The evidence in this chapter 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 as to the extent to which size and productivity are correlated and reallocation enhances productivity.

The evidence shows that countries that open their markets to trade have better static and dynamic allocative efficiency and in turn higher productivity. The covariance between size and productivity rises in response to trade reform and the evidence also shows that market selection improves with trade reform. By the latter, we mean that less-productive businesses are more likely to exit and more-productive businesses are more likely to survive. This improved market selection contributes positively and substantially to productivity growth.

While theory and evidence provide support for trade reform in terms of improved allocative efficiency and associated increases in productivity, both theory and evidence also point towards many things that can go wrong that either mitigate or potentially limit the gains from trade reform. In a highly distorted economy there are second-best problems which mean that piecemeal trade 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, labour and financial markets. A poorly functioning labour 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, which can be very costly in a poorly functioning labour market. 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. Many of these new 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 start-ups 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 and 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 spread globally during crises. Financial regulation that helps monitor the health of the financial services industry and provides safeguards against financial collapses is 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 start-ups 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 the effects of caution and waiting. 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.

Endnotes


2. Of course it is also of interest to ask what induces such large differences in productivity and profitability across businesses in the same sector. As discussed earlier in the introduction, the evidence suggests this partly reflects idiosyncratic choices of product quality and mix, location of the business, organizational practices and the like. It also reflects differences in entrepreneurial and managerial ability. In addition, it most likely reflects a form of luck – being in the right place at the right time with a product and process that is of high value and can be produced in a cost-effective manner. In what follows, as a shorthand we mostly refer to all these factors as differences in productivity (broadly defined) across businesses.
3. See Bernard et al. (2007) for an excellent review of the evidence on firm heterogeneity and trade.

4. 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 (for example, distribution facilities). Both establishment- and firm-level evidence is relevant. For job reallocation, the establishment level is preferred since the frictions in the labour 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, analysing 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 do not formally model multi-establishment firms.

5. Foster et al. (2008) examine 11 detailed product classes for the United States where direct measurement of physical output and prices is feasible. They find that the dispersion of physical productivity is slightly larger than the dispersion of revenue productivity (essentially price times physical productivity). Interestingly, the reason is that physical productivity and price are inversely correlated at the establishment level. This latter pattern is consistent with models of product differentiation such as those in Melitz (2003) and Melitz and Ottaviano (2008).

6. Taking into account both the cross-sectional variation in productivity and size and the dynamics of productivity of size. For example, a recent interesting paper by Giovanni and Levchenko (2010) argues that the very skewed size distribution mitigates the impact of trade reform on aggregate outcomes because even if trade reform impacts market selection as in Melitz (2003) and Melitz and Ottaviano (2008), the large, mature firms that dominate aggregate outcomes are not much impacted. Taking into account the very skewed distribution of firm size is clearly important but this work neglects firm dynamics so that large, mature firms are essentially not subject to productivity shocks (other than perhaps a random exit shock). As discussed in this section (and at some length in Haltiwanger et al., 2010b) even large, mature firms experience a high pace of reallocation and, given this, it is important to make sure such churning enhances productivity.

7. This statistic is consistent with the evidence in Foster et al. (2008).

8. Although Bartelsman et al. (2009a,b) caution against simple cross-country comparisons of the contribution of reallocation to productivity growth. The reasons include measurement and conceptual problems. A better approach is to find some way to explore differences-in-differences that exploit both within-country and between-country variation. That is, suppose that some sectors in a country face more onerous misallocation distortions – then one would expect that it is in those sectors that we observe reallocation to play less of a productivity-enhancing role within a given country than sectors with less onerous distortions.

9. See Davis et al. (2010b) and references therein.


11. Bartelsman et al. (2009a,b) provide evidence on differences across countries on a wide range of distortions.
12. There may be an ameliorating effect on the duration of unemployment 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 that it reflects workers and firms in the informal sector as being underemployed for the reasons discussed in the text.

13. 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.

References


—; —; —. 2010b. "Labor market flows in the cross section and over time", mimeo.


Pagés, C. 2010. The age of productivity: Transforming economies from the bottom up (Inter-American Development Bank, Palgrave MacMillan).


