Almost two decades ago, trade economists began to study the effects of openness when product markets are imperfectly competitive. The resulting new trade literature emphasized that openness increases welfare by increasing the menu of products available in each country. It also emphasized that openness can make product markets more competitive, and can thus induce firms to exploit scale economies more fully. The former message has remained central to modern trade theory, but the view that openness leads to significant gains in firm-level scale efficiency is no longer widely held.

There are several reasons that the link between trade and scale efficiency has been deemphasized. One is that trade economists have found new ways to link openness with welfare. But equally important, the early empirical evidence for trade-induced gains in scale efficiency was less than compelling. Enthusiasm for large scale efficiency effects was initially stoked by simulation studies that found commercial policy reforms might generate efficiency gains on the order of 5 percent or larger (Tybout 1993). However, these simulations ignored intraindustry heterogeneity. Thus, all firms within an industry were treated as being of average size for that industry, and since most industries are populated by many small firms and a few large ones, this average size was typically inefficiently small. This meant that modest increases in firm size could generate substantial efficiency gains, even though most production really came from plants well above minimum efficient scale. Another problem with many simulation models was that they were hardwired to ensure that increases in the scale of production took place at all firms with trade liberalization. But as econometric studies linking import competition and firm size emerged, it became clear that plants in import-competing industries tend to contract when exposed to heightened competition from abroad.

The chapter by Baldwin and Gu (hereafter BG) is interesting because it goes some way toward resuscitating the notion that openness might generate significant welfare gains through simple scale effects. Because it focuses on the length of production runs rather than firm size, scale efficiency gains are possible at both small and large firms. Thus, BG’s model emphasizes a kind of scale effect that can be reconciled with the fact that large firms account for most output. Similarly, it reconciles reductions in the aggregate scale of output with increasing production-run-level scale efficiency.

Perhaps most importantly, BG show that changes in trade policy are sig-
nificantly correlated with product diversity, with production run length, and with firms’ relative positions in the size distribution. Thus, they convincingly demonstrate that trade-induced production run effects are worth worrying about. Their findings complement earlier studies that established a correlation between openness and firm-level efficiency but were unable to control for production run length. They also complement recent work by Bernard, Redding, and Schott (2006b), who show that two-thirds of manufacturing plants in the United States alter their mix of 5-digit products over a five-year period, and this reallocation process appears to be efficiency-enhancing.

The theory developed by BG also constitutes a useful contribution, and nicely complements some recent works on trade, heterogeneous firms, and endogenous scope. Perhaps the one closest to BG is Bernard, Redding, and Schott’s (2006a, hereafter BRS), which characterizes each firm’s productivity in each product as dependent on both firm-level ability and firm-product-level expertise. Higher firm-level ability raises a firm’s productivity across all products, which induces a positive correlation between a firm’s intensive (output per product) and extensive (number of products) margins. The BRS model differs from BG’s because (a) it assumes Dixit-Stiglitz preferences and thus fixes firms’ mark-ups parametrically, and (b) it adds an additional dimension of firm heterogeneity. Nonetheless, its predictions are quite similar: “Trade liberalization fosters productivity growth within and across firms and in aggregate by inducing firms to shed marginally productive products and forcing the lowest-productivity firms to exit. Though exporters produce a smaller range of products after liberalization, they increase the share of products sold abroad as well as exports per product” (BRS, 2006a, abstract).

Also relevant is Nocke and Yeaple’s (2006) model. It too allows for firm heterogeneity and endogenous scope. However, unlike BRS and BG, Nocke and Yeaple assume that span of control problems cause marginal production costs to rise as additional product varieties are added at a given firm. Nonetheless, as in BRS and BG, reductions in trade costs cause the most efficient firms—that is, those with greatest managerial ability—to shed product lines.

The evidence reported in BG and BRS goes some way toward establishing that firms’ scope and production run lengths affect efficiency and are related to trade. But there is much more to explore, and these papers will hopefully inspire further research. One unresolved issue is the role of global fragmentation of production in driving scope. When firms can have a stage of their production done abroad, does this mean they produce less product varieties at home, and to what extent does this phenomenon explain the reduction in firms’ scope observed in the Canadian data? A second issue is how unilateral changes in trade policies affect patterns of scale, scope, and efficiency. As Melitz and Ottaviano (2008), BRS, and Nocke
and Yeaple (2006) have shown, model predictions can depend critically upon the balance between domestic firms’ access to foreign markets, and foreign firms’ access to domestic markets. Third, it would be worthwhile to explore the sensitivity of firm-level responses to trade policy to both the time horizons involved and the entry costs that new firms face. In the short term, or when entry/exit barriers are prohibitive, all of the response to a trade policy change must be accomplished through adjustment by incumbent firms. But over longer time horizons, entry and exit can dampen pressure to adjust on remaining firms. Finally, the sheer magnitude of the tariff effects documented in BG is puzzling, and merits further investigation. Although tariffs only changed a few percentage points on average, they are quite significant in many of the regressions BG report.  

References


3. Similar large effects appear in other studies of the Canada–U.S. FTA (e.g., Head and Reis 1999 and Trefler 2005).