Firm Heterogeneity, Endogenous Entry, and the Business Cycle: Comment*

Paolo Pesenti
Federal Reserve Bank of New York, NBER and CEPR

August 2011

Conventional analyses of the role played by technology shocks over a business cycle typically focus on the adjustment of prices and quantities by a given, constant set of homogeneous firms. Keeping things simple, in the absence of nominal rigidities a rise in productivity allows firms to charge lower prices and increase production under full employment of productive inputs. Productivity increases can have short-term contractionary effects in the labor market if prices lack flexibility, as firms unable to lower prices find it optimal to reduce temporarily their demand for inputs. Either way, the whole adjustment occurs along the intensive margin, without entry or exit of productive units, without creation or destruction of output varieties, without changes in sectoral composition.

This paper similarly studies the propagation of technology shocks in the macroeconomy. But the focus here is on the entry and exit of heterogeneous firms and the process of intrasectoral reallocation, that is, on the extensive

*Thanks to Mary Amiti and Aysegul Sahin for useful comments and suggestions. The views expressed here are those of the author, and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve System, or any other institution with which the author is affiliated.

margin of adjustment. An exogenous increase in labor productivity creates incentives for new firms to enter the market. To the extent that the new (net) entrants are less efficient than the incumbent ones, average productivity falls, dampening the positive effects of the technology shocks on aggregate output per worker. Also, as demand elasticities and market shares change, productive resources move away from more efficient firms, once again dampening the effects of supply shocks on aggregate output.

The bulk of the literature on the extensive margin of adjustment — in both closed and open economies — concentrates on the static, long-run, steady-state responses of labor and product markets to permanent shocks (say, the effects of trade liberalization). Less explored — at least in relative terms — are the higher-frequency dimensions, in particular the role of entry and exit as a propagation mechanism of technology shocks. This paper's attempt to bridge "macro" and "micro" dimensions of analysis and bring them together within a coherent, integrated framework is certainly welcome, its modeling strategy is promising, and the interplay between the static features of endogenous entry and the more traditional dynamic apparatus imported from the real business cycle tradition is rich.

Perhaps too rich. The model introduces many features which are not fully activated or explored in depth. There are no simulations or calibrations, let alone econometric estimates. The numerical analysis toward the end of the paper is useful but does not feed an appetite for a thorough empirical investigation of how endogenous entry impacts cyclical fluctuations in aggregate productivity, markups and prices.

The model extends Melitz and Ottaviano (2008). In both contributions there is firm heterogeneity — modeled in terms of productivity differences as in Melitz (2003) — and endogenous markups, derived through a linear demand system with horizontal product differentiation. At the technical

level, the previous paper considered quasi-linear utility (constant marginal utility of income) while the new model introduces variable marginal utility, a refinement aimed at improving the model's ability to match empirical stylized facts. Last but not least, the previous contribution focused on market size and trade liberalization, while the new paper zooms in on shocks to labor productivity.

For a brief overview of how the model works, it is useful to focus on a simplified (static) analysis of the key building blocks.

In the model, N firms enter the product market by making an irreversible investment. Supply of a product variety requires a variable input (labor) whose requirements per unit of output, indexed by c, are firm-specific. Firms get to know their efficiency c (the lower the better) only once they enter the market. They produce only if they are sufficiently efficient (that is, if the realization of c is below some threshold $\bar{c}$). The (truncated) cumulative distribution function for c is $G(c) = (c/\bar{c})^k$. Firms hire labor and pay a wage $W = 1$ (the numeraire of the model is the efficiency unit of labor). The marginal cost is therefore $Wc = c$.

Firms face a linear demand for their products:

$$q = \frac{\lambda L}{\gamma} (\bar{p} - p)$$

where λ is the marginal utility of income, L is the number of workers (and the market size), γ is a measure of product differentiation, $\bar{p}$ is the choke price (products with price above $\bar{p}$ are not purchased). The profit maximization problem can thus be written as:

$$\max_p \frac{\lambda L}{\gamma} (\bar{p} - p)(p - c)$$

Solving, and observing that $\bar{p} = \bar{c} \geq c$, we conclude that more productive firms (with lower c) have bigger output

$$q(c) = \frac{\lambda L}{2\gamma} (\bar{c} - c),$$

have higher markups

$$p(c) - c = \frac{\bar{c} - c}{2},$$

and yet quote lower prices:

$$p(c) = \frac{\bar{c} + c}{2}.$$

Aggregating, we can write average output Q/N as:

$$\frac{Q}{N} \equiv \frac{\int_0^{\bar{c}} q(c) dG(c)}{N} = \frac{\lambda L}{2\gamma (k+1)} \bar{c}. \tag{1}$$

Each of the L workers supplies z units of labor, so z is aggregate labor productivity. Total wage incomes are therefore:

$$WLz = N \int_0^{\bar{c}} cq(c) dG(c)$$

After some algebra the previous expression yields:

$$\frac{N}{L} = \frac{2\gamma (k+1)(k+2)}{\lambda L k} \frac{z}{\bar{c}^2}. \tag{2}$$

To obtain an expression for average output per worker Q/L, we observe that:

$$\frac{Q}{L} = \frac{Q}{N} * \frac{N}{L} \tag{3}$$

Using (1) and (2) with (3), we derive average output per worker as a positive function of labor productivity z and a negative function of the cutoff $\bar{c}$, with an elasticity that depends on the degree of firm heterogeneity k:

$$\frac{Q}{L} = \frac{k+2}{k} \frac{z}{\bar{c}}$$

The paper shows that when z increases, the cutoff $\bar{c}$ increases as well, reducing the expansionary effects of higher labor productivity. Summarizing, the basic story is that higher labor productivity induces entry (and survival)

by relatively less productive firms. This process is associated with higher prices and markups, and smaller output per firm. But even if there is no entry, there are effects on average output per worker through the marginal utility of income. If the demand elasticity is variable, during an upswing it falls more for high-price firms than for low-price firms. For a given number of active firms, market shares are reallocated toward less efficient firms, so that the benefits from higher labor productivity are dampened by pro-cyclical entry of, and reallocation toward, less efficient firms. Similar considerations hold in an open-economy environment.

The key premise of the paper is that these reallocation effects are relevant at business cycle frequency. Based on what we know about adjustment along the extensive margin (nicely documented among others in a series of contributions by Bartelsman, Haltiwanger and Scarpetta 2004, 2009a, b) we have reasons to expect that the mechanisms emphasized in the paper are empirically relevant, albeit with some nuances.

Needless to say, there is overwhelming evidence on significant firm heterogeneity, large disparities in firm size, firm growth, and productivity performance across markets and countries. The model emphasizes that creative destruction and net entry are important for promoting productivity growth, and the evidence supports this view. However, the contribution of net entry to productivity growth is stronger in high-technology industries compared with low-technology ones, even though differences between these two groups vary significantly across countries. Regrettably, sectoral differentiation is not captured in the model and left to future research efforts.

Successful entrants tend to expand rapidly. In fact, surviving firms are not only relatively larger but also tend to grow rapidly (not an aspect captured in the paper). Perhaps more than net entry, the evidence is that the continuous process of restructuring and upgrading by incumbents is essential to boost

aggregate productivity. Again, this stylized fact plays little role in the current version of the model.

Moving to open economies, the trade literature has developed important research agendas on the effects of trade liberalization on aggregate productivity, with much less emphasis on cyclical effects. For instance, a recent study by Gopinath and Neiman (2011) on the effects of the Argentinean crisis concludes that shocks that raise the costs of production impact aggregate measured productivity through three main channels. First, each individual firm cuts back on imports of intermediate inputs. The size of impact depends on the elasticity of substitution among and between domestic and foreign varieties. Second, standard national accounting practices estimate real imports using a price index which ignores changes in varieties. Failure to account for variety adjustment at the firm level impacts measured productivity. Third, there is a reallocation of market shares across firms that impacts aggregate TFP measurements. Interestingly, this latter element is explicitly emphasized in the model through the role of variable marginal utility.

In conclusion, the paper seems well equipped to deal with many important aspects of adjustment at the extensive margin and its effects on aggregate productivity. The next step is to bring the model to the quantification stage. How big is the dampening effect associated with reallocation and net entry? As the high-frequency implications of the extensive margin adjustment are still relatively unknown, the value added of this applied research agenda is high.

References

[1] Bartelsman, E. J., J. Haltiwanger and S. Scarpetta, 2004, "Microeconomic Evidence of Creative Destruction in Industrial and Developing

Countries." World Bank Policy Research Working Paper No.3464, December.

[2] Bartelsman, E. J., J. Haltiwanger and S. Scarpetta, 2009a, "Measuring and Analyzing Cross-Country Differences in Firm Dynamics", in T. Dunne, J. B. Jensen and M. J. Roberts, eds., *Producer Dynamics: New Evidence from Micro Data*, Chicago: University of Chicago Press, pp.15-76.

[3] Bartelsman, E. J., J. Haltiwanger and S. Scarpetta, 2009b, "Cross-Country Differences in Productivity: The Role of Allocation and Selection." National Bureau of Economic Research Working Paper No.15490, November.

[4] Gopinath, G. and B. Neiman, 2011, "Trade Adjustment and Productivity in Large Crises." National Bureau of Economic Research Working Paper No.16958, April.

[5] Melitz, M., 2003, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica* 71, pp.1695-1725.

[6] Melitz, M., and G. Ottaviano, 2008, "Market Size, Trade and Productivity," *Review of Economic Studies* 75, pp.295-316.