DIEGO RESTUCCIA
University of Toronto

# Comment

An essential issue in economics is understanding why some countries are rich and others poor. A consensus view has emerged in the literature whereby productivity is at the core of the differences in income across nations. Over the last fifteen years, progress has been made in our understanding of cross-country income differences in part by the increasing recognition of the importance of production heterogeneity—firms—and the allocation of factors of production across them for aggregate outcomes. The progress has been enhanced by the much wider availability of microeconomic data sets of firms across a growing number of countries. There has been a productive interplay of macro- and microeconomic approaches to development. Two papers in this issue of *Economía* reflect very well the synergy that is growing across subfields, and both papers provide valuable insights for the overall role of firms and productivity on development. Although both papers offer insights and implications that are broad across many fields in economics, I focus my discussion in the context of the macroeconomic development literature and, in particular, the connection of misallocation and aggregate productivity.

In “Firm Dynamics and Productivity: TFPQ, TFPR, and Demand-Side Factors,” John Haltiwanger discusses relevant measurement issues surrounding estimates of firm-level productivity and the implied inference of distortions, misallocation, and aggregate productivity derived from a variety of microeconomic data sets of firms.

What is typically meant by misallocation? The concept of misallocation is tightly related to a particular economic structure. To focus the discussion, consider a simple one-period economy where a single homogeneous good is produced by heterogeneous establishments that differ only in their productivity. That is, more productive establishments produce more output for a given set of inputs. Assume for illustration that there are decreasing returns to scale in inputs at the establishment level, although this assumption is not essential. Incidentally, the aggregate production function implied by this setting features constant returns to scale with the number of firms as an input, consistent with standard practice in macroeconomics. In this simple context, the efficient allocation of inputs across establishments is one that maximizes total output given an aggregate set of inputs. The efficient allocation involves assigning more inputs to more productive establishments and the same amount of inputs to equally productive establishments, so that all establishments feature the same physical marginal product of factors. More productive establishments are larger than less productive ones precisely to achieve equalization of marginal products. Misallocation occurs when an allocation deviates from the efficient benchmark allocation, although it may well be that the efficient allocation can never be achieved in practice. Moreover, the aggregate cost of misallocation can be expressed as the ratio of actual aggregate output to efficient output as determined by the efficient allocation. With this setup in mind, it is easy to see how market and centralized economies can generate allocations that differ from the first best. Examples include any tax or subsidy to a subset of establishments, even if well intended to promote productive enterprises or sectors; any allocation of inputs not guided by productivity, regardless of whether this is done via price distortions or centralized arrangements with limitations to reallocation; and any market friction that prevents the allocation of resources to the best uses. It is also easy to see that both the measurement of productivity at the microeconomic level, in this case the establishment, and the measure of misallocation critically depend on the specific structure/abstraction, such as one good, one sector, no dynamics. As a result, any relevant misspecification will translate into a biased picture of productivity and misallocation. The abstraction focuses on a subset of potential sources of misallocation, but in practice misallocation can occur in many other settings, such as across producers in different sectors, across occupations, across space, across ethnic or racial groups, or across time; and the importance of these sources can differ across time and space.

As Haltiwanger points out, in most productivity studies, microeconomic data provide the value of outputs (or more generally revenue) for a given establishment or firm. Hence, the issue discussed at length is how to recover physical measures of productivity at the firm level and how inferences about deviations of allocations to these measures of productivity depend on particular assumptions about the market structure. Specifically, the most influential empirical application of misallocation was provided by Hsieh and Klenow, who, using microeconomic data for manufacturing plants in China, India, and the United States, document the presence of misallocation in these countries and the large reallocation gains to the efficient allocation on aggregate productivity in China and India relative to the gains in the United States.[1] From the perspective of Haltiwanger's analysis, Hsieh and Klenow's calculations are plagued by all the measurement and specification issues raised in the paper. In fact, there are probably other relevant issues as well. An unproductive feature of the literature that has followed Hsieh and Klenow's approach in other countries and contexts, including Haltiwanger's paper, is the focus on dispersion in TFPR as a summary measure of the extent of misallocation and in some cases its aggregate consequences, as it is clear that a given dispersion in TFPR need not have the same consequences for aggregate productivity—for example, in the case of economies with very different dispersion in physical productivity. These differences will be reflected in the reallocation gains, but not in the dispersion of revenue productivity, as can be seen from the results in Hsieh and Klenow for the different countries and different time periods.[2] It is also clear that for measurement and specification issues to matter for the main conclusions from Hsieh and Klenow's analysis, the issues need to be systematically more prevalent in poor countries relative to rich countries. The key question, then, is what to do with these difficulties. Measuring productivity at the microeconomic level is a daunting task, but not one that can ever be solved with more and better data. More detailed and better data would certainly help, but at a macroeconomic level it is hard to imagine that these issues can be convincingly solved with just data. There is also not the "right" model, as any model is false by construction; instead a model is relevant to the extent that it represents a reasonable abstraction for the question at hand. The more productive approach, as suggested by Haltiwanger, is to explore the extent to which different assumptions affect the nature and quantitative magnitude of the inferences from data and theory.

From a macroeconomic perspective, while the empirical evidence of misallocation, such as that in Hsieh and Klenow and the large follow-up studies, has been extremely useful in motivating a large body of inquiry into the nature and consequences of misallocation, the empirical methodology is ultimately limited because of the lack of connection of misallocation to the specific policies and institutions that create it. As a result, I re-emphasize the view expressed in my earlier work with Rogerson that "the most persuasive evidence in support of the role of misallocation will come from work that follows the direct approach in specific contexts, especially those in which we

1. Hsieh and Klenow (2009).
2. Hsieh and Klenow (2009).

observe changes in some underlying source of misallocation and can measure the resulting change in misallocation and aggregate TFP."[3] That is, the literature should move beyond documenting the extent of misallocation and instead study specific policies and institutions as potential sources of misallocation and exploit changes in policies and institutions to assess their quantitative impact. When confronted with the myriad of measurement and specification issues involved in documenting misallocation, production heterogeneity and misallocation may not seem as important. But the opposite perspective arises with regard to the actual specific policies, frictions, and institutions that create misallocation. For example, the caste system in India discriminates in education and occupations based on class as opposed to ability; the hukou system in China imposes severe restrictions on labor migration across space and sectors and prevents the reallocation of land to the most productive uses in the agricultural sector; the presence of state-owned enterprises in many developing countries disrupts efficient market allocation; and the role of financial institutions in providing and channeling credit affects the incentives to enter and exit markets. The issue is how important these policies and institutions are quantitatively. A recent literature explores the linkage of these distortions in the context of models in which the productivity of microeconomic units is endogenous to the economic environment, an issue emphasized by Haltiwanger. I elaborate more on the connection of the productivity distribution and misallocation in my discussion of Hopenhayn's paper below.

Haltiwanger discusses evidence from data on firm-level prices to draw some implications of the connection of distortions to measures of revenue productivity in the United States and the potential implications across countries. A useful addition to this discussion from the macroeconomic perspective is the evidence from microeconomic data in agriculture, where actual quantities for outputs and inputs are reported. In the case of poor countries, agriculture is the sector where most economic activity occurs, and agricultural data have the advantage that output is easily recorded in standard physical units, so the analysis is less subject to measurement and specification errors. However, much of the work in development economics focuses on the role of frictions in farm growth and within household allocations. In contrast, Adamopoulos and Restuccia emphasize the role of misallocation in understanding the low aggregate productivity in the agricultural sector in poor countries.[4] Restuccia

3. Restuccia and Rogerson (2013).
4. Adamopoulos and Restuccia (2014).

and Santaeulalia-Llopis use detailed microeconomic data from Malawi, a very poor country in Africa, that corroborates the findings of the productivity literature of wide dispersion in productivity across production units, in this case household farms, including within those farms that mostly produce a single crop: maize.[5] Because most of the land in Malawi is customary without well-defined property rights, households in Malawi are given use rights of the land within a local community, and these use rights are distributed among the members on a fairly egalitarian basis. Restuccia and Santaeulalia-Llopis's analysis corroborates the findings of the importance of misallocation in the context of severe restrictions to land markets. For instance, a subset of farms with access to marketed land are shown to be much closer to their efficient allocation than farms with no marketed land, providing a direct connection to the role of land markets.

Within the agricultural literature, there are also studies focusing on particular reforms over time. For instance, Kirwan, Uchida, and White study the Tobacco Transition Act of 2004 in the United States, an abrupt policy liberalization of production restrictions.[6] The paper estimates that of the substantial gains in productivity that followed the reform, all are accounted for by the reallocation of resources across productive farms and the productivity impact of entry and exit. Adamopoulos and Restuccia study the role of reallocation after a policy reform—namely, the land reform in the Philippines that put a cap on the maximum land size of farms—using panel data of farms before and after the reform.[7] The results show a large negative productivity impact of the reform, especially the government interference in the land market and targeted allocation of land to landless and smallholder farmers.

To summarize, there are many important measurement and specification issues at the core of misallocation and productivity analyses. As emphasized by Haltiwanger, a productive area for future work is the study of specific policies and institutions (or changes over time) where detailed quality data may be available and where the abstraction required (assumptions on technologies and preferences) may be more appropriately targeted. Also relevant is the recognition that the productivity distribution of microeconomic units may be responsive to and interact with many of the policies and institutions emphasized.

5. Restuccia and Santaeulalia-Llopis (2015).
6. Kirwan, Uchida, and White (2012).
7. Adamopoulos and Restuccia (2015).

In “Firm Size and Development,” Hugo Hopenhayn studies the relationship between average establishment size and development observed in a cross section of countries. The paper examines this relationship through the lens of a standard model, such as the models proposed by Lucas and Hopenhayn.[8] The starting point is the documentation of a strong positive relationship between average establishment size and GDP per capita across countries. The importance of this empirical finding should not be overlooked, as there is quite a bit of confusion about the exact relationship in the empirical literature. The confusion arises from the sparse and often inconclusive evidence on average establishment size across a large set of countries. The challenges include data availability that typically restricts the analysis to a handful of countries (or larger samples with mostly richer countries); the reporting of data that disproportionately include larger establishments; and data that include multiple sectors in the economy with potential differences in sector-specific operational scales. For these reasons, Bento and Restuccia construct a newly assembled data set for the average employment size of manufacturing establishments across a large sample of countries using census or representative survey data.[9] The finding, as reported by Hopenhayn for a set of Latin American countries, is a strong positive relationship between average establishment size and GDP per capita. The average establishment size in the top decile of the GDP per capita distribution is nineteen workers (persons engaged), versus six in the bottom decile, that is, a more than threefold factor difference between the top and bottom deciles of the income distribution. In the United States, the average establishment size is twenty-two workers, whereas in India it is three.[10] As emphasized by Bento and Restuccia, failing to account for unpaid and self-employed workers or smaller establishments severely biases the average establishment size not only in the United States, but especially in poorer countries where the vast majority of workers are engaged in small or own enterprises. This explains the disparate results in previous studies.

Having established the main empirical fact of interest, which is that the average establishment size in rich countries is substantially larger than in poor and developing countries, Hopenhayn turns to the question of what can account for this fact. Hopenhayn considers a model of production

8. Lucas (1978); Hopenhayn (1992).

9. Bento and Restuccia (2016).

10. For reference, in Bento and Restuccia’s sample of 134 countries, the difference in GDP per capita between the top and bottom deciles is forty-six-fold.

heterogeneity that encompasses simpler versions of the model developed by Lucas and Hopenhayn.[11] The theory developed generates several interesting insights into the relationship between average establishment size and development. The paper shows how variations in the number of firms, the productivity distribution of firms, and returns to scale generate a positive correlation between average establishment size and GDP per capita. However, a quantitative version of the model is able to generate large variations in average establishment size with arguably extreme variations in parameters and small variations in GDP per capita. In a nutshell, the variables considered generate a much larger income elasticity of establishment size than observed empirically.

There are two points worth emphasizing in the context of Hopenhayn's framework and results. First, exogenous differences in total factor productivity (TFP) across countries have no effect on average establishment size. Imagine that the establishment production function in Hopenhayn is altered so that the labor productivity of the establishment is $Az$, where $z$ is idiosyncratic productivity and $A$ is a common residual TFP component across all establishments. Because entry and operational costs are all scaled up with wages and hence output, a reduction in $A$ in poor countries has no implications for average establishment size and only reduces GDP per capita relative to rich countries. Similarly, this result implies that growth in TFP over time for individual countries would have no effect on establishment size. This implication is consistent with the setup of the Lucas model used by Guner, Ventura, and Xu and the version of the Hopenhayn model in Bento and Restuccia.[12] This is an important implication of the framework for at least two reasons: (a) the evidence from developed countries is that within sectors, establishment size has not changed much even though there is roughly constant positive productivity growth, which lends confidence in the framework to study differences across countries; and (b) rich and poor countries differ in many dimensions that can be encapsulated in residual TFP ($A$) differences such as human capital and technology adoption, but these differences are not direct competitors to the channels emphasized in Hopenhayn as potential sources of establishment size differences across countries.

Second, policies and institutions that create misallocation and lower aggregate productivity have either no effect on establishment size or the same effect

11. Lucas (1978); Hopenhayn (1992).
12. Guner, Ventura, and Xu (2008); Bento and Restuccia (2016).

created via changes in the number of establishments analyzed in Hopenhayn. For instance, in the version of the Hopenhayn model used by Bento and Restuccia, if there is no investment in productivity, misallocation has no effect on the equilibrium number of establishments (as would be the case in a version of the Hsieh-Klenow model with endogenous entry).[13] Similarly to the previous point on exogenous residual TFP differences, misallocation would act as a level shifter and with scaled costs implies no change in the number of establishments. This suggests that for misallocation to be a potential source of establishment-size differences, some specific features are needed.

There is a large literature emphasizing the potential impact of misallocation on aggregate productivity. The literature discusses at length the type of policies and institutions that create misallocation, so I do not dwell on that in this comment. But an important development within this literature involves attempts to endogenize the productivity level of individual establishments—that is, unlike Restuccia and Rogerson, for instance, where the productivity distribution is exogenous and constant across countries, the recent literature makes the productivity of establishments respond to the same policies and institutions that create misallocation.[14] I provided an early example of this approach in an application to productivity in Latin America.[15] Other examples, in addition to the references mentioned in Hopenhayn, include Bhattacharya, Guner, and Ventura and Da-Rocha, Tavares, and Restuccia.[16] Within this literature, the papers by Bento and Restuccia and by Hsieh and Klenow are tightly connected to the average establishment size observations across countries reported by Hopenhayn.[17] The key insight behind these papers is that policies and institutions that systematically discourage productive establishments, what Restuccia and Rogerson call correlated idiosyncratic distortions, have the potential not only to create misallocation as emphasized in the earlier literature, but also to discourage investment in productivity either on entry or over the life cycle.[18] Many policies and institutions take this form even if not intended by design: a general tax may only be applied to or enforced on large firms; an export tax affects more productive firms that export; and financial frictions effectively affect productive, larger-scale or growing enterprises. Importantly, this property of distortions is what is found empirically by Hsieh and

13. Bento and Restuccia (2016); Hsieh and Klenow (2009).
14. Restuccia and Rogerson (2008).
15. Restuccia (2013).
16. Bhattacharya, Guner, and Ventura (2013); Da-Rocha, Tavares, and Restuccia (2016).
17. Bento and Restuccia (2016); Hsieh and Klenow (2014).
18. Restuccia and Rogerson (2008).

Klenow for India, Mexico, and the United States and by Bento and Restuccia for a larger set of countries.[19] The effect of correlated distortions on productivity investment is intuitive: distortions that disproportionately affect productive establishments act as a tax on the return to productivity investment and hence have the potential to lower productivity on entry and productivity growth over the life cycle.

How important are these mechanisms for generating differences in average establishment size and GDP per capita across countries? The analysis of Bento and Restuccia reveals that correlated distortions generate sizable effects both on establishment size and aggregate productivity, but the channels through which this occurs were not well understood.[20] An important feature of their analysis is the contrast of the effects of correlated distortions on entry-level productivity with the effect on life-cycle productivity growth emphasized by Hsieh and Klenow.[21] A relevant observation emphasized in Hsieh and Klenow's paper is that the productivity gap between young and old establishments grows much wider in developed countries such as the United States than in developing countries such as India. This empirical fact is very appealing as a potential candidate for lower establishment size and aggregate productivity in poor countries. However, while it is important to understand the economic determinants of this fact, the findings of Bento and Restuccia suggest that the dynamics of firm productivity are less useful for thinking about income and establishment size differences across countries. To appreciate that differential life-cycle growth is unlikely to generate large amplification of income differences, note that while lower life-cycle growth implies lower aggregate productivity in a cross section of establishments, it implies less dispersion in productivity across establishments, such that distortions have less of a negative effect on aggregate productivity. In a calibrated version of a Hopenhayn-type model with life-cycle investments, Bento and Restuccia show that these two opposing effects on aggregate productivity roughly cancel each other, implying that life-cycle growth differences have no substantial amplification effect on aggregate productivity.[22] Hsieh and Klenow find a similar quantitative result in a more detailed model of life-cycle investment.[23] In addition, life-cycle growth differences may not generate large differences in establishment size if distortions substantially reduce

19. Hsieh and Klenow (2014); Bento and Restuccia (2016).
20. Bento and Restuccia (2016).
21. Hsieh and Klenow (2014).
22. Bento and Restuccia (2016).
23. Hsieh and Klenow (2014).

the dispersion in establishments' productivity by age. This is the case in Hsieh and Klenow, where correlated distortions reduce productivity growth in the life cycle but more so for productive establishments. As a result, correlated distortions generate a relatively small reduction (about 15 percent) in average establishment size in India in their study. Bento and Restuccia find a larger life-cycle effect of correlated distortions on establishment size because their setup is such that Gibrat's law holds even in the distorted economy, so distortions affect the life-cycle growth of all establishments in the same proportion.

To be sure, it is unlikely that a single source is responsible for the patterns of establishment size across all countries, as in some countries entry costs may be more severe than correlated distortions or the other channels analyzed in Hopenhayn; but overall the results in Hopenhayn are very useful because they indicate the likely channels through which policies and institutions play a role in accounting for the observed operational scale differences across countries. As Hopenhayn emphasizes, more work is needed in identifying and assessing the quantitative relevance of specific policies and institutions that are at the core of establishment size and aggregate productivity differences across countries.

## References

Adamopoulos, Tasso, and Diego Restuccia. 2014. "The Size Distribution of Farms and International Productivity Differences." *American Economic Review* 104(6): 1667–97.

———. 2015. "Land Reform and Productivity: A Quantitative Analysis with Micro Data." Working Paper 540. University of Toronto, Department of Economics.

Bento, Pedro, and Diego Restuccia. 2016. "Misallocation, Establishment Size, and Productivity." Working Paper 557. University of Toronto, Department of Economics.

Bhattacharya, Dhritiman, Nezih Guner, and Gustavo Ventura. 2013. "Distortions, Endogenous Managerial Skills, and Productivity Differences." *Review of Economic Dynamics* 16(1): 11–25.

Da-Rocha, José-María, Marina Mendes Tavares, and Diego Restuccia. 2016. "Firing Costs, Misallocation, and Aggregate Productivity." Working Paper 561. University of Toronto, Department of Economics.

Guner, Nezih, Gustavo Ventura, and Yi Xu. 2008. "Macroeconomic Implications of Size-Dependent Policies." *Review of Economic Dynamics* 11(4): 721–44.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long-Run Equilibrium." *Econometrica* 60(5): 1127–50.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124(4): 1403–48.

———. 2014. "The Life Cycle of Plants in India and Mexico." *Quarterly Journal of Economics* 129(3): 1035–84.

Kirwan, Barrett E., Shinsuke Uchida, and T. Kirk White. 2012. "Aggregate and Farm-Level Productivity Growth in Tobacco: Before and after the Quota Buyout." *American Journal of Agricultural Economics* 94(4): 838–53.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9(2): 508–23.

Restuccia, Diego. 2013. "The Latin American Development Problem: An Interpretation." *Economía* 13(2): 69–100.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics* 11(4): 707–20.

———. 2013. "Misallocation and Productivity." *Review of Economic Dynamics* 16(1): 1–10.

Restuccia, Diego, and Raul Santaeulalia-Llopis. 2015. "Land Misallocation and Productivity." Working Paper 533. University of Toronto, Department of Economics.