AGRICULTURAL PRODUCTIVITY AND STRUCTURAL TRANSFORMATION: EVIDENCE AND QUESTIONS FOR AFRICAN DEVELOPMENT

Douglas Gollin
Agricultural Productivity and Structural Transformation: Evidence and Questions for African Development*

Douglas Gollin
Oxford University†
April 1, 2021

Abstract

This paper summarises key findings from the recent literature on agricultural productivity and structural transformation and then identifies priority areas for further research. The paper discusses the macro relevance of agricultural productivity for growth and structural transformation. New theory and data have underscored the importance of agricultural productivity as a proximate cause for low aggregate productivity in the world’s poorest countries, including many in sub-Saharan Africa. Evidence has also emerged on the importance of agricultural productivity growth as a driver of development and growth in the past. But micro evidence and changes in the global context suggest that today’s low-income countries may face a different set of challenges in the years ahead.

---

*This paper was prepared as a Pathfinding Paper for the CEPR-led research programme on Structural Transformation and Economic Growth (STEG), funded by the United Kingdom’s Foreign, Commonwealth and Development Office (FCDO) as part of the UK aid effort. Portions of this paper were previously provided as a briefing note for the Agricultural Global Practice team of the World Bank. This paper has benefited from my prior work and conversations with Stefan Dercon, Cheryl Doss, Remi Jedwab, Tim Kelley, David Lagakos, Richard Rogerson, James Stevenson, Chris Udry, Dietz Vollrath, Mike Waugh, and many others; none of them bear any responsibility, however, for what I have written. The views expressed in this paper are entirely my own and do not reflect the official viewpoint of any organisations for which I work or with which I am affiliated.

†Contact: Douglas Gollin, Oxford Department of International Development, Oxford University, 3 Mansfield Road, Queen Elizabeth House, Oxford OX1 3TB. E-mail: douglas.gollin@qeh.ox.ac.uk
1 Introduction

Across sub-Saharan Africa, over 60% of the population lives in rural areas, and agriculture remains the dominant source of employment in most countries of the region (World Bank, World Development Indicators; henceforth WDI). However, measured productivity levels are extremely low. In a proximate sense, as pointed out by Caselli (2005) and Restuccia et al. (2008), these two facts alone imply an unpleasant agricultural arithmetic for Africa: if many people earn their living from agriculture, and if agricultural productivity is low, then aggregate incomes will be correspondingly low. But this problem is compounded by a sectoral gap in productivity: value added per worker in African agriculture appears to be far lower than comparably measured labour productivity in other sectors, even after adjusting for differences in input quantity and quality Gollin et al. (2014).

What accounts for the low aggregate productivity of African agriculture? Indeed, is it as low as it appears, or does mismeasurement explain a large part of the apparent sectoral productivity gaps? To the extent that productivity gaps are real, are they evidence of any inefficiency? Would increases in agricultural productivity serve to drive growth and economic transformation? If so, under what conditions? When, where, and how would agricultural productivity growth lead to sustained, sustainable, and inclusive growth? What obstacles, frictions, and barriers limit productivity growth within agriculture? What are the implications for policies directed at agriculture or rural development?

This paper addresses these and other questions in assessing the role of agricultural productivity in achieving structural transformation and growth in sub-Saharan Africa. The paper summarises a range of recent literature, both micro and macro in focus, and it looks to sketch out some of the directions for current and future research. The focus is explicitly on countries in sub-Saharan Africa, but much of the same analysis applies to heavily rural and agricultural economies in Asia and Latin America, or at least to the more isolated rural regions within those economies.

The past fifteen years of scholarship have substantially advanced our understanding of the relationship between agricultural productivity growth and broader economic transformation and development. An important insight has been the heterogeneity of this relationship, even
within countries. Almost by definition, rural areas face spatial frictions that limit their access to markets. Agricultural commodities are typically heavy and bulky; many are also perishable. This means that spatial heterogeneity is of first-order importance in thinking about the pathways for agricultural productivity improvements to stimulate growth and transformation. Productivity improvements also have different implications depending on the physical and economic characteristics of different crops and commodities. Taken together, these two dimensions of heterogeneity suggest that we should not assume that there is a single relationship between agricultural productivity growth and any particular outcome. In some cases, agricultural productivity growth can lead to overall growth and the expansion of non-agricultural activities; in other cases, it can lead to specialisation in agriculture. In some cases, agricultural productivity growth might drive inclusive growth and poverty reduction; in other contexts, it may deliver growth alongside the displacement of the rural poor. This heterogeneity poses a challenge for research: average effects, at the macro level, may not be greatly informative. But as research on agriculture and growth has accumulated in recent years, the broad outlines of the complex relationship have begun to emerge.

The remainder of this paper is organised as follows. Section 2 will summarise recent evidence, both macro and micro, on the centrality of agricultural productivity in understanding income and productivity differences across countries, and on the sectoral gaps within countries. This section will also discuss the ways in which the literature has wrestled with difficult issues of measurement and selection. Section 3 addresses theories related to agriculture’s role in the growth process, and the ways in which recent evidence has challenged these theories. Section 4 summarises the literature on market failures and misallocation in agriculture. Section 5 focuses on agricultural technology and innovation. Section 6 speculates on the prospects for replicating the Green Revolution experience of the late 20th century. Section 6 discusses directions for further research and implications for development strategy. Section 7 concludes.
2 Sectoral Gaps and Cross-Country Productivity Differences

The importance of agricultural productivity differences in explaining cross-country income differences was the focus of work by Córdoba and Ripoll (2009); Gollin et al. (2007); Restuccia et al. (2008); Vollrath (2009), and others. These papers underscored that, in a proximate sense, cross-country income differences are substantially explained by differences in agricultural productivity. Given that, the productivity gaps between agriculture and other sectors should be a prominent focus for growth research.

2.1 Evidence on agricultural productivity gaps

In most developing countries, the share of the workforce in agriculture substantially exceeds the share of GDP that is derived from agriculture. (See Figure 1). To the extent that both measures are accurately constructed, the arithmetic implication is that value added per worker (referred to by economists as ‘average labour productivity’) is lower in agriculture than in other sectors of the economy. Commonly used national aggregate data suggest that labour productivity is around four times higher in non-agriculture than in agriculture, a very large difference. Although there is no formal reason why one should expect labour productivity to be equalised across sectors, equality is a useful benchmark. Relative to that benchmark, the gaps are quite substantial.

Gollin et al. (2014) show that this ‘agricultural productivity gap’ (APG) remains, although it shrinks considerably, when agricultural output is measured carefully and labour use in agriculture is adjusted for differences in hours worked as well as a number of observable dimensions of human capital per worker, as shown in Figure 2. These adjustments reflect the fact that workers in agriculture generally have fewer years of schooling than those in other sectors; and their schooling may in fact be of lower quality than that received by workers in other sectors.

There are a number of plausible interpretations of the APG. Different interpretations lead in different directions, in terms of the implications for our understanding of agriculture’s role in growth and development. One view is that the APG reflects a problem of misallocation.
If we took the average product of labour as a proxy for the marginal product, not an altogether unreasonable assumption with a standard constant-returns production technology, then the large APGs in developing countries would tend to suggest that labour should flow out of agriculture into other sectors. With diminishing returns, this process should eventually lead to equalisation of productivity across sectors. Rural-urban migration does seem to be a ubiquitous demographic feature of low-income economies. But to the extent that the APGs seem to have persisted in many countries for long periods of time, some frictions or wedges could be preventing productivity from equalising across sectors. But other views are also possible. APGs could be caused by differences in capital shares across sectors, or by differences in the quality of labour along dimensions unrelated to education. The next sections explore these explanations in greater detail.
2.2 Measurement issues

One interpretation of the APG is that effective labour units are poorly measured, even after the adjustments made in Gollin et al. (2014). For instance, McCullough (2017) examines household survey data from sub-Saharan countries (rather than the census and labour force survey data used in Gollin et al., 2014) and finds that agricultural workers spend many fewer hours working than workers in other sectors. This raises the possibility that the APG would be greatly reduced if labour were measured more accurately. Arthi et al. (2018) find evidence that commonly used survey designs tend to overstate the hours worked by farmers. They find that with more frequent interviews, and correspondingly shorter recall periods, farmers report fewer hours worked. This suggests that even the household survey data may still be overstating hours worked in agriculture.

To the extent that better measurement lowers or eliminates the APG, it suggests that labour markets work well. In this view, the APG reflects a ‘gap’ in average productivity but does not reflect any obvious friction or distortion. In this view, the agricultural sector might continue to shrink in size as the economy grows (consistent with many theories of structural change), but there would be no reason for policies to hasten the movement of workers out of agriculture. The same conclusion would hold under another interpretation, consistent with efficiency, under
which the APG could reflect systematic differences in the capital-intensity of production across sectors. In the presence of such differences in capital intensity, the equalisation of marginal products of labour across sectors would be consistent with large differences in average products of labour.

2.3 Sorting and selection

Another possible explanation for APGs is that there could be significant and systematic differences in the quality of workers across sectors, based on unobservables such as ability or motivation. In this case, the apparent differences in labour productivity across sectors would simply reflect mismeasurement of the labour inputs. If agriculture attracts lower ability individuals, then the APG could simply arise from a tendency to overstate the effective units of labour in agriculture. This view has been supported by studies (e.g. Hicks et al., 2017; Herrendorf and Schoellman, 2018) suggesting that for individuals who switch sectors, moving from agriculture to non-agricultural employment, wages do not rise substantially. If these ‘switchers’ do not increase their earnings substantially, then labour markets might be working effectively, supporting the idea that the APG reflects differences in the skill composition of the workforces in the two sectors (as argued in Young, 2013), rather than a distortion in labour markets.

2.4 Evidence on living standards and migration

An alternative interpretation of the APG is that it might stem from frictions that limit the mobility of labour and capital across sectors. If individuals cannot move freely to sectors with higher marginal products of labour (and if capital is also limited in its free movement across sectors), the natural tendency towards equalisation would be stymied. In support of this view, many studies find that individuals who move from rural areas to urban areas experience substantial increases in real earnings (e.g. Bryan and Morten, 2019; Munshi and Rosenzweig, 2016), and some careful randomised controlled trials (e.g. Bryan et al., 2014) show what appear to be important frictions in labour markets. As noted above, though, other studies challenge the idea that real earnings differ significantly for individuals who transition across sectors or locations (Hicks et al., 2017; Herrendorf and Schoellman, 2018).
Comparisons of earnings and wages are inevitably difficult, because of the need to adjust for differences in the composition of consumption baskets, on the one hand, and prices, on the other hand. One approach is to look at real measures of consumption or living standards. This is the approach taken by Young (2013), who constructs an index based on measures of real living conditions; he finds large differences between rural and urban averages in living standards. In very much the same vein, Gollin et al. (2017) consider a number of measures corresponding to real living conditions, across locations within 20 sub-Saharan countries. A clear pattern is that measures of household well-being improve almost monotonically with population densities within countries. Perhaps more surprising, the same holds for public goods provision and for a number of types of amenities, including crime and exposure to air pollution. Although the data do not offer a clear picture of the importance of selection and sorting on these outcomes, Gollin et al. (2017) conduct a number of tests that challenge the notion that these spatial disparities in living standards are entirely the result of sorting.

Certainly, data on rural-urban migration, which remains substantial in many sub-Saharan countries, tend to support the idea that differences in living standards are both real and highly salient. Across sub-Saharan Africa, people are indeed moving to cities. The process of urbanisation is rapid, and by some accounts, the fastest in human history. In 1950, the urban population of sub-Saharan Africa was approximately 20 million. Fifty years later, in 2000, the urban population had risen by an order of magnitude to over 200 million, and by 2050, this population is projected to grow to 1.26 billion people. United Nations projections suggest that the share of people in sub-Saharan Africa living in urban areas will rise from 41.4% in 2020 to 58.1% in 2050, an average annual growth rate of 4.28% sustained for a century (United Nations, Department of Economic and Social Affairs, Population Division (2018). World Urbanization Prospects: The 2018 Revision, Online Edition). It would be difficult not to view this as evidence that rural-urban gaps, and by extension sectoral gaps, are real, rather than artifacts of measurement.

2.5 Interpreting the APG

The interpretation of APGs has profound implications for policies involving agriculture, rural development, urbanisation, and even industrial policy. If the APGs emerge from market failures,
then there may be a role for government policy to shape the structural transformation of the economy, whether by reducing frictions or by pursuing more activist policies around job creation in non-agriculture. If the APGs instead shrink or disappear under closer scrutiny and tighter measurement, then it may suggest that labour markets are working well. In this case, faster intersectoral movement of labour is not a goal in itself, and we should expect that structural change will take place organically via changes in productivity or shifts in trade costs and market access.

There are numerous needs for research to shed further light on the nature of sectoral productivity differences. Better measurement can clarify the extent to which the APGs are real and the extent to which they are driven by selection on observed or unobserved skills (as in Lagakos and Waugh, 2013). Another topic for research might be to examine in detail the extent of labour market integration between rural and urban areas. How strong is the evidence for frictions in the labour market? What are the sources of these frictions?

3 The Food Problem and Agriculture-Led Strategies for Growth and Transformation

What does economic theory have to tell us about agriculture’s role, and specifically the role of agricultural productivity increases, in structural transformation and growth? There are two predominant strands of theory, based on conflicting stories and yielding different policy implications. In one strand, agricultural productivity growth is a necessary condition for poor countries to begin the growth process. The other strand is far more sceptical and tends to view agriculture as a lagging sector, pulled along by growth that arises in manufacturing or some other sector that displays rapid productivity growth. This section briefly reviews these two strands of thought, both of which can be traced back to the earliest thinking about economic growth.
3.1 Early models, classical and neoclassical

Although it is possible to find much older antecedents for today’s debates over agriculture’s role in development, a more useful starting point is in the early years of modern development economics: i.e. in the 1950s and early 1960s. At this historical moment, with many developing countries attaining independence and embarking on their own development paths, there were two competing visions. In the view of W. Arthur Lewis (1951), or of Ranis and Fei (1961), traditional agriculture in low-income economies was relegated to the ‘subsistence sector’ and was portrayed as a sector with little dynamism that could nevertheless function as a holding pen or reservoir for ‘surplus labour’. In Lewis’s vision, most smallholder agriculture was reproductive rather than productive. Although plantation agriculture and some other commercialised farms belonged in the ‘modern’ sector, Lewis saw the majority of small farms as having an abundance of workers with near-zero marginal products of labour. Long-run growth, for Lewis, consisted of moving workers out of subsistence into the modern sector, where their labour could be “fructified” by capital. In his view, the subsistence sector was almost definitionally incapable of generating any sustained increase in per capita income; although he didn’t fully spell out his argument, he seems to have envisioned a quasi-Malthusian equilibrium holding in the subsistence sector.

By contrast, the other strand of literature argued that agricultural productivity growth was a necessary condition for economies to grow. The economist Theodore Schultz, who ultimately shared a Nobel Prize with Arthur Lewis, wrote that many poor countries suffer from a “food problem” (1953). He defined this as the need for countries to produce the vast majority of their own food, before they can begin to move resources into other sectors. Schultz believed that “transforming traditional agriculture” (to use the title of his most widely read book) would require giving farmers new technology and/or new skills. Schultz’s argument was taken up and formalized by Johnston and Mellor (1961), among others.¹ This view has persisted for several decades in the world of agricultural economics and agricultural development, not least in the works of Mellor (1973; 1995; 2017) and Timmer (1988; 1992; 2002).

¹ Another early work, that of Nurkse (1953), offered a view that somehow combined both views of agriculture. Nurkse identified the agricultural sector of poor countries as a source of under-utilised labour that could be released at near-zero marginal cost and used for the construction of public and private capital goods with high rates of return. In this sense, his view echoed that of Lewis (1953).
3.2 Contemporary closed-economy models

In the past fifteen years, a new literature has emerged seeking to model structural transformation within the context of standard growth models. Most of the literature has followed Solow (1956) and Swan (1956) by working with closed economy models, typically extending the Solow-Swan framework by incorporating multiple sectors. The simplest way to model a transformation is to use a simple two-sector growth model with a minimum consumption requirement for agriculture. Within this class of models, one way to generate a structural transformation is with some combination of exogenous technological change, delivering Solow-type growth, and non-homothetic preferences, which lead consumers to shift the composition of their purchases as incomes rise. In particular, if the income elasticity of demand for agricultural goods is below one, then the budget share of agricultural goods will fall as the economy grows. Since the economy is assumed to be closed, this necessarily leads the agricultural sector to shrink as a share of value added and the labour force. This approach was central to papers such as Goodfriend and McDermott (1995), Echevarria (1997) and Kongsamut et al. (2001). In all of these models, growth in agricultural productivity helps economies to move out of agriculture, but so does any other kind of income-enhancing change in the economy.

A variant of these closed economy models was put forward in Gollin et al. (2002) and Gollin et al. (2007). In these papers, a minimum consumption requirement for agricultural goods takes on the stark form that consumers derive utility only from food consumption, up to a level of satiation, beyond which they derive no incremental utility from expanded food consumption. In these models, technological progress in agriculture is the key to initiating the process of modern economic growth. The “food problem” keeps poor economies from achieving any meaningful growth, but once they embark on the process of industrialisation, exogenous technical progress continues as in the standard one-sector growth model, so that these economies converge asymptotically to a one-sector Solow model. The model also gives a striking implication, which is that economies that are bad at producing food will take a very long time to begin the growth process. This framework is useful for thinking about environments in which agricultural productivity is fundamental to economic growth, but the framework also highlights the dependence of the results on key assumptions. In particular, the model depends
crucially on the closed-economy assumption. For poor countries, trapped in low productivity agriculture, imported food would allow them to jump into non-agricultural production, where they can benefit from robust exogenous growth.

Non-homothetic preferences provide one mechanism for structural transformation in a growth model. But other mechanisms are also possible. Another potential way in which sectoral transformation can occur is through differential productivity growth, still in the context of a closed economy. With multiple sectors experiencing productivity growth at different rates, what happens to the size of those sectors? As is obvious on reflection, the answer to this question depends on the interplay between the differential growth rates and the question of whether the goods produced in the two sectors are net substitutes or net complements. If they are net substitutes, then the faster-growing sector will effectively displace the slower-growing sector: the economy does not particularly need the slow-growing sector, because a substitute is available at a decreasing relative price. If they are net complements, however, then the faster-growing sector will actually shrink, as more resources move into the slower-growing sector. In effect, capital and labour inputs will move into the slow-growing sector to compensate for the relatively slow pace of productivity growth, because this economy continues to need the good that is produced in this sector. The latter phenomenon, of a slow-growing sector expanding in size relative to the rest of the economy, is a variant of Baumol’s cost disease story (originally sketched out in Baumol and Bowen, 1966). A more complex, more general version provides the framework in Ngai and Pissarides (2007). Using this framework, agriculture might shrink with economic growth if (a) it is the slow-growing sector, and agricultural goods are broadly substitutable with non-agricultural goods and services; or (b) agriculture is the fast-growing sector, and the goods are complementary. Arguably the Ngai-Pissarides story offers a better account of the growth of services in rich countries than of the decline in agriculture’s importance in low-income countries.

Note that these two mechanisms for generating structural transformation in a closed economy setup are not mutually exclusive. It is possible for both to operate simultaneously. Identifying which mechanism is at work in the data is difficult; Alvarez-Cuadrado and Poschke (2011) attempt to tease apart the two mechanisms using historical data on prices. Both frameworks rely on the assumption of a closed economy, in which countries must produce the goods that
they consume (and vice versa). If this assumption is relaxed, then in a sense consumer demand becomes irrelevant. In a frictionless open economy, consumption and production are divorced. Countries specialise in production of those goods in which they have a comparative advantage, and they consume those goods that they prefer and can afford, with no regard to the country of origin.

3.3 Openness and the missing food problem

Challenging the closed economy assumption leads to a different set of puzzles. Almost all countries in the world produce a substantial share of their own food. Even in rich countries, where agriculture is a relatively small fraction of the economy, and where productivity levels are high, domestic production of food and agricultural goods is typically substantial in relation to consumption. To some extent this reflects the sheer weight and bulk of agricultural commodities, as mentioned above; and to some extent it also reflects local tastes and preferences. What accounts for the low levels of trade in food? Why do poor countries import so little food (and other agricultural goods) when they seem to be so unproductive, in both absolute and relative terms?

Tombe (2015) tackles this question using a model with labour market distortions and trade costs. In this model, international trade frictions, presumably some combination of policy distortions and transport costs, keep poor countries from importing the efficient quantities of agricultural goods, and these countries must produce the remainder of their needs from domestic farms. In related work, Teignier (2018) finds that trade in food played an important role in the industrialisation processes of Great Britain and South Korea at the relevant points in their historical development. The magnitude of this effect depends strongly on the other ingredients in the model; Świecki (2017) considers a model with multiple frictions and barriers to structural transformation and finds that international trade frictions account for a relatively small share of the cross-country differences in the extent of structural transformation.
3.4 Matsuyama and long-term growth

In a closed economy with a subsistence requirement for agricultural goods, it is beneficial to low-income countries to have high agricultural productivity. But what are the implications for open economies? Matsuyama (1992) considers an environment where the non-agricultural sector is characterised by a learning-by-doing technology, such that the sector displays a kind of dynamic increasing returns. In autarky, countries benefit from having high agricultural productivity, since it allows them to move more resources into non-agriculture, which then leads to earlier and faster growth. However, in an open economy setting, countries with high relative productivity in agriculture end up specialising (partly or completely) in agricultural production, so that they grow more slowly than economies that are less productive in agriculture. This implies that a country with a rich endowment of arable land could find its natural abundance to be a mixed blessing. High productivity and output in the agricultural sector may, without the offsetting changes in relative prices that would occur in a closed economy, squeeze out the manufacturing sector.

Matsuyama shows that a country whose relative productivity in agriculture is greater than for the rest of the world will specialise in agriculture and will thus grow slower than the rest of the world. Initial patterns of comparative advantage will also intensify, since the rest of the world becomes increasingly better at non-agriculture. Eventually, the agricultural economy may specialise completely.

The welfare implications of this model are not entirely clear. The agricultural economy does experience an improvement in the terms of trade as the rest of the world becomes more productive. There are clearly welfare gains for the agricultural economy from raising productivity levels. Conditional on a country having a comparative advantage in producing agricultural goods, it is better off having a greater absolute advantage. However, for an economy with approximately the same relative productivity as the rest of the world, there might be a benefit from tipping its comparative advantage into non-agriculture rather than agriculture.
3.5 Testing the models: evidence from empirical micro literature

Testing models of structural transformation is necessarily difficult. Since we cannot easily or ethically carry out experiments on entire economies, researchers have a limited set of tools. A number of the papers referenced above rely on quantitative exercises in which the models are taken to data, in a variety of ways. The behaviour of the model economies can be compared to the data from actual economies. To avoid overfitting, the models are ideally compared to data ‘out of sample’ (i.e. different from those that were used to calibrate them). Sometimes, macro models are used to motivate cross-section or comparative analysis of different economies, but in macro economic settings, it is normally difficult to find convincing ways to distinguish causal relationships from correlations.

An alternative approach that has provided valuable insights in recent years is the use of careful micro-scale studies that address structural transformation and growth in within-country contexts where data quality is potentially higher and causal identification is feasible. A number of papers have specifically asked how agricultural productivity growth affects the cross-sector movement of labour and economic activity. Most of these papers have taken advantage of some plausibly exogenous geographic within-country variation in agricultural productivity and have then sought to link that variation to a range of outcomes that are informative about structural transformation.

(Bustos et al., 2016) examine the impact of two agricultural technologies that arrived in Brazil in the 1990s: genetically modified soybeans and double cropping of maize. These technologies affected different locations within Brazil, and the technologies also had different characteristics. The new soy technologies were labour-saving, while the maize technology was labour-using. (Bustos et al., 2016) trace the differential impacts of the two technologies on non-agricultural labour in Brazilian municipalities. They argue that, possibly because of the differing characteristics of the two technologies, the local impacts on sectoral movements of labour varied. Since both technologies were productivity enhancing, a useful and important message is that we should not expect agricultural productivity growth to deliver the same impact on structural transformation and growth across all contexts. As argued in (Dercon and Gollin, 2014), we should expect substantial heterogeneity in this pathway, rather than looking
for a single theory that will hold in all contexts and at all moments in time.

These findings are consistent with the somewhat ambiguous story that emerges about a different agricultural technology in the US. Hornbeck and Keskin (2014) consider the local economic impacts of an irrigation technology that benefited certain locations in the Midwest of the US but not others, with the variation due to the (exogenous) location of a vast but irregularly shaped underwater aquifer. Hornbeck and Keskin (2014) found little or no evidence of major economic impacts, including structural change, from the diffusion of the new technology in the 20th century.

Somewhat in the same vein, Asher and Novosad (2020) examine the local labour market impact of rural road construction in India, using a fuzzy regression discontinuity design that takes advantage of comparisons across towns and villages of similar size and characteristics, some of which received roads through a national construction programme. Rural roads should, in principle, increase agricultural productivity by increasing the farmgate prices of output and reducing the farmgate costs of purchased inputs; this improvement in terms of trade should be essentially homologous to an increase in the physical productivity of agriculture. However, the authors find only modest impacts on the local economies of villages that received improved roads. The largest impact seems to have been a modest shift in employment, with some workers moving from agricultural labour into manual labour in nearby towns.

Among the other papers that have looked at local labour market impacts of agricultural productivity shocks, typically short-term weather shocks, are papers by Emerick (2018), Santangelo (2016), and Colmer (Forthcoming, 2021). These papers also show how changes in agricultural productivity can spill into the rest of the economy, albeit on a much more short-term basis. These papers may be picking up rather different mechanisms than the longer-run shifts in the sectoral composition of output. For instance, they may be capturing short-term intersectoral flexibility of markets, very much linked to population density, rather than longer-term structural shifts. Nevertheless, the availability of high-quality micro data allows these studies to capture causal effects quite cleanly.
3.6 Spatial frictions and their importance

An emerging area of literature in economics has focused on within-country heterogeneity and spatial frictions, in relation to agricultural productivity and economic specialisation. Using spatially explicit models and a new toolkit of general equilibrium analysis, these papers have allowed researchers to ask questions about the role of domestic spatial frictions in determining the geography of agricultural production and the movement of labour across sectors. The issue of spatial frictions is particularly acute for many developing countries, including especially those in sub-Saharan Africa, which have few roads (and poor quality roads) and which therefore have large populations in quasi-subsistence modes of production. Traditional trade models have focused on trade frictions between countries and have assumed that within-country frictions are relatively unimportant. The new literature, however, challenges this view. Papers model spatial frictions within countries, e.g. Atkin and Donaldson (2015), Costinot and Donaldson (2016), Costinot et al. (2016), Donaldson (2018), Porteous (2016), and Sotelo (2018). To what extent do these frictions alter crop choice, affect input use, and prevent local specialisation? Do these frictions alter the direction or pace of structural transformation? This has emerged as an exciting and promising area for research.

4 Misallocation and Gains from Liberalisation

How else can we explain low productivity? One idea that has received considerable attention in recent years is that there are simply too many farmers. The sheer number translates into many small farms; perhaps these are unable to take advantage of scale economies or lumpy capital investments. But another explanation has intrigued researchers in recent years. With more than half the adults in the region earning their living from agriculture, it seems plausible that not all are equally capable.

We expect there to be a distribution of farming skill, reflecting knowledge and cognitive abilities, physical strength, motivation, and a myriad of other qualities. With well-functioning markets, economists would expect the least effective farmers to move out of agriculture into other occupations, either selling or renting their land to farmers who are more skillful or energetic. This process should winnow the pool of farmers, leaving the most skilled. But farmland
is not widely transacted in sub-Saharan Africa. In many countries, communities retain control of land, with village chiefs or elders parceling land out to families based on their perceived needs and abilities. This may achieve some broad egalitarian objective, but perhaps the lack of markets is one reason for low productivity.

In short, a simple explanation for low agricultural productivity could be that lots of farms are operated by relatively low-skill farmers. This explanation is appealing, not least because it matches some evidence and observations about productivity in manufacturing firms in the developing world. Hsieh and Klenow (2009) provided evidence that the most productive manufacturing firms in China or India operate on a par with the most productive firms in the US, but average productivity is reduced in China, and even more in India, by the large number of relatively unproductive firms. This research raised the possibility that poor countries may have institutional frictions or rigidities that prevent unproductive firms from exiting the market. The result is a kind of aggregate inefficiency, resulting from the misallocation of labour, capital, and managerial effort.

Could this story hold for agriculture in Africa? Work by Adamopoulos and Restuccia (2014), Adamopoulos et al. (2017), Chen et al. (2017), Restuccia and Santaeulalia-Llopis (2017) and Deininger et al. (2017), among others, has pursued this analysis. Research has benefited from the availability of detailed and nationally representative datasets on farming, such as the World Bank’s Living Standards Measurement Surveys and their Integrated Surveys of Agriculture. These have allowed researchers to construct careful measures of total factor productivity (TFP) at the farm level for a number of countries in sub-Saharan Africa (and elsewhere). A consistent finding is that there appears to be huge TFP dispersion across farms within a given country.

It is certainly possible that, at the level of individual firms or production establishments, misallocation of inputs can reduce productivity. Among manufacturing firms, researchers have found clear evidence of poor management practices (Bloom and Van Reenen, 2007; Bloom et al., 2010, 2013), and there is also strong evidence that misallocation may arise from inefficient capital markets (Banerjee and Duflo, 2005; De Mel et al., 2008; Karlan and Morduch, 2010; Karlan et al., 2012). It does not seem implausible that inefficient firms and farms may survive in an environment where productive resources are allocated through networks of patronage and family ties, rather than through competitive markets. It also seems possible that the inefficiency
generated in this way could create a consequential drag on aggregate productivity. Restuccia and Santaeulalia-Llopis (2017) estimate that the misallocation created by land market frictions in Malawi could reduce agricultural output by as much as two thirds. (Gollin and Udry, 2021) argue that this analysis tends to confuse misallocation with other sources of productivity dispersion, including weather and pest shocks, measurement error of various kinds, and heterogeneity in the quality of land and other inputs. Nevertheless, they find that misallocation could lead to substantial losses in aggregate efficiency in agriculture, and hence to sizable reductions in overall output.

Given the pervasive market failures that characterise agriculture in poor countries, it would be unrealistic and unconvincing to argue that misallocation is not an important hindrance to growth and structural transformation. Not only does misallocation across farms keep productivity low in agriculture, but there may also be frictions that prevent efficient allocation of resources across sectors. For instance, Gottlieb and Grobovšek (2019) argue that frictions in land markets prevent households from exiting agriculture to the extent that would be efficient. A similar story is at the heart of Chen (2017). Whether it would be sufficient for growth to remove or reduce these frictions is less clear. Are these misallocations really binding constraints? Or do they instead pose problems that high-functioning economies would find ways to solve? Is misallocation a cause of low growth and transformation, or is it simply a feature of poor economies that are experiencing limited growth?

5 Improved Technology and Its Impacts

One of the key pathways of increased agricultural productivity has been the introduction of new agricultural technologies that increase TFP. The role of new agricultural technologies in inducing structural transformation and economic growth has long been a subject of interest for historians and students of agricultural development. Scholars of the industrial revolution in Britain have argued at length and in detail about the centrality of agricultural productivity growth to the industrial revolution. For instance, Nurkse (1953) wrote:

Everyone knows that the spectacular industrial revolution would not have been possible without the agricultural revolution that preceded it. And what was this
agricultural revolution? It was based mainly on the introduction of the turnip. The lowly turnip made possible a change in crop rotation which did not require much capital, but which brought about a tremendous rise in agricultural productivity. As a result, more food could be grown with much less manpower. Manpower was released for capital construction. The growth of industry would not have been possible without the turnip and other improvements in agriculture.

This view has been widely contested in the economic history literature, in debates far too wide-ranging to summarise here. (For a sample of views, see Allen (1994, 2000, 1999) and Overton (1996).) The role of agriculture in industrialisation in other parts of the world, besides Britain, also suggests a diverse set of experiences and interpretations.

5.1 Historical evidence and interpretation

Many examples from history offer guidance of large-scale effects. There is evidence on the effects of the adoption of potatoes in Europe following the Colombian Exchange (Nunn and Qian, 2011). Easterly (2007) looks at the impact of plantation agriculture on long-run development patterns, and Bentzen et al. (2017) investigate the impact of irrigated agriculture on political outcomes.

5.2 Macro evidence from the Green Revolution

Perhaps the strongest and most powerful evidence for agriculture-led growth relate to more recent times and to a literature examining the experience of the Green Revolution. From its beginnings in the 1960s, the Green Revolution attracted the attention of economists who hoped to understand the processes of growth and evolution. The challenge is that understanding the causal impacts of Green Revolution technology is not straightforward. Agricultural productivity growth is not randomly distributed across countries, and it is almost never independent of growth in other sectors. As noted above, it is easier to find tight causal identification at the micro level or the regional level, using natural experiments or quasi-experiments, as in Bustos et al. (2016) and Hornbeck and Keskin (2011). Another approach has been to bring structure to the problem (e.g. Foster and Rosenzweig, 2004, 2007). However, even when these localised
effects can be well identified and carefully measured, they may not generalise or scale easily to full general equilibrium impacts on aggregate economies. The local movements of people across sectors are only a relatively small part of the transformation process. In particular, for poor countries with large fractions of their workers initially in agriculture, the main mechanisms of structural transformation are not played out within local labour markets. Instead, they involve large-scale movements of people across locations, from rural to urban or from one region to another. Studies that emphasise the local movements of people will miss these broader and more secular changes.

However, a recent set of papers have used spatial and time variation to look at the impacts of the Green Revolution. These include Moscona (2019), Bharadwaj et al. (2020), der Goltz et al. (2020) and Gollin et al. (Forthcoming, 2021). These papers all emphasise that the nature of the scientific advances was such that the Green Revolution served as an essentially exogenous productivity shock to recipient countries. The papers all use some variant of agro-climatic conditions or previous crop mix, combined with evidence on the development of the Green Revolution technologies in different crops, to instrument for the country-specific and time-specific impacts of the Green Revolution. Gollin et al. (Forthcoming, 2021) use a Bartik-type instrument to see how economic outcomes vary in relation to different countries’ exposure to Green Revolution innovation. These papers all find that the Green Revolution technologies had strong impacts on agricultural productivity. Bharadwaj et al. (2020) and der Goltz et al. (2020) both find a statistically significant link to improvements in infant mortality. Gollin et al. (Forthcoming, 2021) find large, positive, and significant effects of the Green Revolution on income per capita in countries that benefitted from the technologies, and they find some evidence for structural change, with agricultural productivity moving workers out of agriculture into other sectors. Moscona (2019), in contrast, finds evidence that the direction of structural change was, if anything, in the opposite direction, with the Green Revolution pulling workers into agriculture and making recipient locations more agriculture-intensive. An overview of this literature suggests that there is still much to be done in order to understand the historical effects of the Green Revolution, but it is heartening to see that a large non-experimental event of this kind is susceptible to rigorous \textit{ex post} empirical analysis. There is also, of course, a large literature in which the impacts of the Green Revolution have been estimated using models (e.g.
Evenson and Gollin, 2003), but the results of these model-based analysis are hugely sensitive to the structure of the models, and it is difficult to validate the underlying models effectively.

5.3 Micro evidence and experimental findings

A far greater body of literature has built up trying to look at the impacts of agricultural technology on structural change, using micro data of various kinds. Some of the most interesting evidence comes from long-term mixed-method studies of structural change in village communities. The seven-decade-long study of Palanpur, India, is perhaps the most striking of these (Himanshu et al., 2018) as it traces structural transformation within a rural community. Earlier work in the Philippines documented similar transformations in the lives and livelihoods of a set of communities that had been followed for over thirty years (Hayami and Kikuchi, 1999).

Shorter and smaller-scale studies of the impact of agricultural technology have been numerous. Recent surveys in the economics literature include Renkow and Byerlee (2010) and Pingali (2012). This paper is also related to an even larger literature which has considered the impact of agricultural science or research on economic and social outcomes at a more geographically limited scale. This literature has been surveyed by Maredia and Byerlee (2000); other important contributions include Fan et al. (2002), Meinzen-Dick et al. (2003), Thirtle et al. (2003), Pingali and Kelley (2007), Dalrymple (2008), Raitzer and Kelley (2008) and Rusike et al. (2010).

6 Replicating the Green Revolution: Challenges and Questions

To the extent that the Green Revolution may have led to significant economic growth and structural transformation, it immediately raises the question: can we repeat this episode today? Can a new Green Revolution bring prosperity and transformation to those parts of the world that continue to lag in agricultural productivity, and where poverty remains acute? Was the Green Revolution a replicable historical event, or can it be repeated if only the right technological innovations can be found?
6.1 What’s different in the 21st century?

But does this story still apply? As noted above, one lesson from the empirical literature (e.g. Bustos et al., 2016) is that not all technologies lead to the same positive impacts. We know that some productivity-enhancing technologies have the potential to exacerbate poverty, for instance, the introduction of mechanisation that displaces landless agricultural labour. Other technologies may increase production but lead to a crash in prices for crops with relatively inelastic demand. Where consumers are rich, producers are poor, and demand is relatively inelastic, the distributional impacts of increasing productivity may not be desirable. (This might be true, for instance, of the global market for coffee or cocoa.) To the extent that the Green Revolution generated positive impacts on growth and development, we should not assume that future technological innovations will be benign.

More to the point, we should wonder whether today’s developing countries are different from the economies of South and Southeast Asia in the mid 1960s, in ways that may pose challenges for future Green Revolutions. For a start, today’s developing countries are almost all richer and less dependent on agriculture than was the case of the countries that benefited most from the early Green Revolution. With the higher living standards of these countries and greater levels of urbanisation, there are also shifts in diets. Urban diets today are more diverse and less dependent on local supplies than was true of rural diets in Green Revolution Asia. It would be impossible to find a single commodity in the consumption basket of sub-Saharan Africa today that matches the importance of rice and wheat in the countries of Green Revolution Asia.

In this environment, it is difficult to see how we can drive economic growth through staple food technology and production increases. The mechanisms do not seem as strong or robust as they were in the last century. Even within farm households, it is not clear that producers seek yield increases at all costs; they are often interested in reducing labour demand to free up time for off-farm activities. The changing roles of women as producers and consumers also affect the desired technologies.
6.2 What’s different in sub-Saharan Africa?

In sub-Saharan Africa, an additional set of constraints arise. One emerging point of difference, relative to the Green Revolution, is that today’s African cities are far less dependent on their surrounding agricultural hinterlands. Our models of rural-urban linkage have generally assumed that rural areas play a key role in supplying urban areas with food and raw materials. This was to a large extent true in the historical context in Europe and North America, and it also seems like a good description of East Asia, South Asia, and Southeast Asia, especially in the 1960s. However, this model does not offer a particularly good characterisation of sub-Saharan Africa today.

In today’s world, urbanisation and infrastructure improvements put many people in sub-Saharan Africa, and the rest of the developing world, within plausible reach of food imports. In many African countries, this has led to an emerging disconnect between urban consumption and rural production. Africa’s urban consumers appear to be moving rapidly away from the foods that are produced in rural areas, a striking difference from Asia in 1960s, when urban consumers simply wanted larger quantities of grains.

Interesting evidence on this pattern is found in recent work by Cockx et al. (2018), who have examined dietary changes in Tanzania. Drawing on household survey data from Tanzania, they found both that urban diets differ significantly from rural diets and that newcomers to cities adopt urban diets quickly. In data from 2008-09, urban residents consumed fewer calories from maize, cassava, and other starchy foods, compared to rural people. Urban consumers got 28% of their total calories (755 kcal per day relative to total intake of 2,652 kcal per day) from these staple foods. By contrast, rural consumers got 57% of their calories from these sources (1,432 kcal per day, relative to total intake of 2,507 kcal per day). Urban residents consumed more rice, bread, pasta, sugar, sweets, pastries, and snack food, almost all of which were imported.

These changes in diets have many reasons. Not least is the rising value of women’s time in urban areas. Rice and bread are convenience foods, as shown in Senauer et al. (1991). In addition, urban households have less kitchen space and face higher costs of cooking fuels than rural households. More meals are eaten away from home. And starchy staples are seen as low status foods, while prepared and processed foods are typically seen (and advertised) as ‘modern’
and ‘sophisticated.’

The shifts in diets have macro consequences. In the context of sub-Saharan Africa, most of these processed, prepared, and convenience foods are imported. For example, in 2014, Tanzania imported around $1.24 billion of agricultural products (according to COMTRADE data). Over 70% of these imports were of processed and semi-processed items: sugar and sweeteners, beverages, tobacco products, processed grains, and dairy products. Processed and semi-processed foods made up 6.9% of the country’s total imports, and they accounted for 1.65% of GDP. In this context, it is not clear that interventions to drive up the productivity of maize or pigeon pea on Tanzanian farms will lead to economic growth and transformation.

Equally important, as rural-urban linkages weaken, the converse causal chain may also fail: urban growth will not necessarily create strong backward linkages into rural areas. Instead, urban and rural growth may become increasingly disarticulated. Urban growth may not drive rural growth, and vice versa. Agricultural growth may come to depend on external markets for cash crops, rather than on domestic markets for food crops. In this context, the Green Revolution may not be easy to replicate.

6.3 What’s different post-COVID-19?

At the time of this writing, the world is wracked by a global pandemic. Beyond the immediate and tragic health and mortality consequences, the COVID-19 pandemic has upended transnational mobility and disrupted supply chains. Within countries, it has led to unprecedented lockdowns and quarantines. As we look to the future, it is worth asking whether the past relationships between agricultural productivity and structural transformation will remain unchanged. Will global trade in food and agricultural goods be dramatically reduced? If so, this could affect both the exports of developing countries and their imports.

Perhaps a more pressing question relates to pathways for sectoral change. In developing countries that have been counting on trade-led industrialisation, will the opportunities remain for participation in global value chains? Will the potential of long-term social distancing alter production processes and induce automation and mechanisation? Will urbanisation slow, as people opt to remain in less densely populated areas with lower rates of disease transmission?

At the moment, it is far too soon to have concrete evidence on these questions, but these are
surely areas for urgent research. Although it is necessarily difficult to do empirical research on the future, many of the models and theory-driven approaches that have been used in the macro literature should be useful in thinking about potential directions and magnitudes of impact.

7 Implications for Development Strategies and Directions for Research

Our understanding of agricultural productivity growth and its relationship to economic development and transformation is remarkably deep but nevertheless has major gaps. At its heart are a number of unanswered questions. We do not, still, have an adequate explanation of why so many people work in quasi-subsistence agriculture; our best accounts depend on unsatisfying assertions about preferences for rural life, or intangible measures of the support provided through social networks. We do not entirely understand what factors drive the process of transformation. Rapid industrial growth clearly has some potential to pull labour out of agriculture, and this process has arguably been the main pathway for transformation in today’s rich countries (as argued, among others, by Lin, 2012). Whether agricultural productivity growth can provide the initial impetus for this process, and whether it needs to, is unclear.

One lesson that emerges fairly clearly from the literature is that there is substantial heterogeneity in agriculture’s role, both across geographical contexts and across time. The power of agricultural productivity growth to drive transformation depends critically on the nature of demand for agricultural goods. This will vary across space, reflecting (among other things) spatial frictions and openness to trade. It will also vary with the specific commodities that are experiencing productivity growth. To give one example, when a country experiences an increase in the productivity of soybeans in an area with good access to international markets, we should expect that it faces a highly elastic global demand curve. If the world price is sufficiently high, we should expect to see domestic resources pulled into soybean production, leading to an expansion, rather than a contraction of the agriculture share of GDP. The effects on agricultural employment are somewhat ambiguous: it will depend on whether the land was previously being cultivated and, if so, on how the labour intensity of soybean compares with the labour intensity
of the crops that were previously grown. A very different story will hold for a staple food crop with little potential for trade, such as sweet potato. An increase in productivity in sweet potato might plausibly drive down prices and encourage a shift of labour out of agriculture into other sectors.

The important role of heterogeneity should remind us to be wary of the simple models that have long dominated policy discussions. The policy question should not be “agriculture first” versus “industrial policy”. Instead, we should be asking far more nuanced questions about when and where agriculture will play a key role in structural transformation, and which commodities or elements of the agricultural sector might be most likely to generate inclusive and transformative growth.

From research, we clearly need to develop models that recognise the spatial heterogeneity of agriculture, models in which countries can be partly open and partly closed, i.e. in which domestic trade frictions are explicitly modelled. These economies may be open at the border but with limited pass-through of imports. We also need to consider models that better describe a process of structural transformation in which the shift from agriculture to non-agriculture is not automatically equated with a shift from rural to urban. This can better account for some of the contexts in which we see structural transformation taking place without migration, e.g. by rural households sending one or more family members to ‘commute’ to town jobs, as seems to be common in parts of Asia. This process can lead to growth and transformation without consolidation of land holdings, a process observed in both China and Japan.

We need our models also to intersect with the complicated (and non-monotonic) movements of labour from home to market (and from market to home), which have a powerful impact on the demand for food. The strongly gendered nature of home production means that women’s employment in market activities outside the home strongly shapes the demand for processed and prepared food. In some settings, this may imply that sectors such as food processing, food retailing, and food service may have a large impact on structural transformation, perhaps larger than the impact of changes in farm-level productivity. To capture these effects, we need models in which the agricultural sector does not directly produce food, but instead produces raw materials for food production. By considering agriculture as an intermediate good, our models can address questions related to productivity in transport and processing. With a few
noteworthy exceptions, e.g. the work of Reardon (2015), Barrett et al. (2019) and Reardon et al. (2019, 2021), relatively little research has focused on the industrial structure of the food and agriculture sector in developing countries. Food value chains have been modelled rather naively in the literature on agricultural development.

We may also need models that explicitly incorporate cash crop production, i.e. the production of commodities that are not directly consumed in the countries where they are produced, but that compete for land and other inputs with food agriculture. To what extent (and when, where, and how best) can these commodities contribute to growth and transformation? Is there an argument for supporting productivity increases (e.g. through research) in these commodities? Or will the consumer benefits flow entirely to consumers in rich countries, with the least productive farmers in poor countries made worse off, at least in relative terms, and possibly in absolute terms? Richer models are needed in which we can consider these fairly detailed questions about the macro consequences of changes in agriculture and food systems, and yet these models need to be sufficiently simple and transparent that the relationships between model assumptions and results are not obscured.

8 Conclusion

Much work still needs to be done to understand better the sources of cross-country productivity differences in agriculture and the importance of these differences for growth and development. We know many of the ingredients that account for these differences. Technology differences across countries are surely important; even in an era where information can travel freely, agricultural technologies need to be quite specifically targeted to geographic locations. For this reason, agricultural technologies typically need to be tailored to very specific geographies, or alternatively the production environment needs to be controlled to match the technology (as in greenhouse agriculture or intensive animal agriculture). Beyond technology, we know that imperfections in the markets for land, labour, and capital may also be important in explaining cross-country productivity differences. The links between rural and urban economies matter, with physical and social mobility costs playing an important role in accounting for the observed differences across countries.
Beyond these static explanations, we badly need to re-think our models of transformation and growth. At present, our models typically assume frictionless trade and movement between urban and rural economies; this is an assumption that needs to be relaxed. Africa’s urban and rural economies are not always closely connected, so that urban growth has taken place largely disconnected from rural economies. In this environment, with significant frictions, growth in one sector may not effectively propel growth in the other. This lack of market integration gives rise to a case for policy interventions, as in Fajgelbaum and Gaubert (2018) and Kucheryavy et al. (2016).

There is scope and opportunity for new research. Detailed datasets now provide insights into smallholder agriculture in a number of countries of sub-Saharan Africa. New models and tools are also available. The time is ripe for further research that moves beyond the one-size-fits-all models and development strategies that have long shaped the debate over agricultural development.
References


Baumol, William J and William G Bowen, Performing arts, the economic dilemma; a study of problems common to theater, opera, music, and dance, New York: Twentieth Century Fund, 1966.


Maredia, Mywish K. and Derek R. Byerlee, “Efficiency of research investments in the presence of international spillovers: wheat research in developing countries,” *Agricultural Economics of Agricultural Economists*, January 2000, 22 (1).


Mellor, John Williams, Agriculture on the Road to Industrialization, Published for the International Food Policy Research Institute, Johns Hopkins University Press, 1995.


