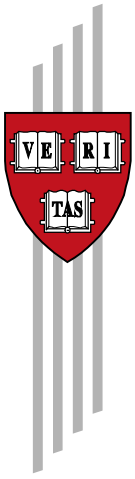
What is South Africa's Crop Production Potential?

Federico Sturzenegger, Bailey Klinger, and Iván Ordóñez

CID Research Fellow and Graduate Student
Working Paper No. 149
May 2023



Working Papers

Center for International Development
at Harvard University

What is South Africa's Crop Production Potential?*

Federico Sturzenegger†, Bailey Klinger‡, and Iván Ordoñez§

May 2, 2023

Abstract

Combining satellite data with FAO potential yields we provide a new measure of South Africa´s current and potential crop farming output. We find that field crop production is twice its census estimate, contributing 1.4% of GDP rather than 0.7%, and that achieving potential could increase its contribution a further 0.5% of GDP. Estimating horticulture potential is more difficult. We find that its 0.7% contribution to GDP is massively unreported, with actual production at 2.5%. Reaching potential could increase this number a further 0.5%. The distance from current to potential output represents over 100 billion 2017 rand of additional gross income and about 350.000 thousand jobs and is unevenly distributed across the country and concentrated in four provinces: Free State, Western Cape, Kwazulu-Natal and Eastern Cape. Our result suggests that there is room to expand agriculture, but because the potential gains are geographically concentrated, the solutions should have a strong location dimension.

*We thank William Beinhart, Peter Delius, André Jooste, Wandile Sihlobo, Johann Kirsten, Ferdi Meyerf and Brian Whitaker for useful conversations and information. The team at CID was fully supportive all through the way in arranging these conversations. A special appreciation is for Thao-Nguyên Bùi, whose help was invaluable to obtain the estimations in this paper. We also thank Patricio Goldstein, Ricardo Hausmann and Tim O´Brien for useful suggestions. Of course all errors are our own. Contact information.

†Harvard Kennedy School, Universidad de San Andrés and HEC, Paris, fsturzenegger@udesa.edu.ar

‡Growth Lab, bwklinger@gmail.com

§Universidad del CEMA, iordonez@gmail.com

1 Introduction

South Africa is an agricultural powerhouse, not only because of its size of 93 million hectares of farmland[1] but because it is embedded in an economy with one of the more developed financial systems in the emerging market world, and is sprinkled with world class agricultural firms with global reach. Yet, there are apparent dimensions where land remains under-utilized. The sector is highly dualistic, and one can find highly productive private commercial farms existing next to communally-administered agricultural land that is low-productivity or even abandoned (Beinart and Delius (2018)). The historical sources of this dualism are clear (Kirsten and Sihlobo 2021), though solutions are not. Much focus has been given to land reform, with an estimated 17% of formerly white-owned land having been redistributed to the state and black-owned farmers (Kirsten and Sihlobo 2021), and arguments are made for further redistribution of commercial land (Cousins 2018). In turn, in a somewhat vicious cycle, uncertainty on the land reform process may have increased perceived risk in agricultural investments and slowed innovation in the commercial sector.

Of course the discussion on policies to improve agricultural production in South Africa has been intense. It is just a question of checking out the (NAMC 2022) to understand the wealth of initiatives. Infrastructure, know-how, financing, property rights all blend into a very large set of policy initiatives, most of which require fiscal resources. Kirsten and Sihlobo (2021) list a string of initiatives discussed over the years. However, the problems seems today as present as ever.[2]

[1] See Kirsten and Sihlobo 2021

[2] These include, the White Paper on Agricultural Policy, 1995, The Strauss Commission Report into the Provision of Rural Financial Services, 1996, the Broadening of Access to Agriculture Thrust, 1996,the Integrated Food Security and Nutrition Programme, 1996, the Integrated Sustainable Rural Development Strategy, 2000, the Land Reform for Agricultural Development Programme, 2001, the Strategic Plan for South African Agriculture, 2001, the comprehensive Agricultural Support Programme, 2004, the Micro Agricultural Financial Institutions of South Africa initiative, 2004, the Proactive Land Acquisition Strategy, 2006,the agricultural and land chapters of the Accelerated and Shared Growth Initiative for South Africa, 2007, the recapitalisation and Development Programme, 2010, the agricultural and land chapters of the National Development Plan, 2011, the

We believe these initiatives should be informed by the answer to a number of relevant questions. First and foremost we need to know how large is South Africa's agricultural potential and how far we are from that potential. It is impossible to assess the benefits of a policy without an estimate of its potential gains. Then there is the question of where in South Africa's should resources to help agricultural production be deployed? Should they focus on commercial agriculture? Should policies focus on the homelands? Should they concentrate in land redistribution? In improving the technology of subsistence farmers? Should it focus in regions with the highest potential gains? If so, which are these areas? None of these questions can be answered without an estimate of the geographical distribution of potential gains.

In this paper we undertake the task of attempting to answer these questions which means calculating the potential output for South African crop agriculture. Our computation combines satellite data, census data, and FAO's estimates of potential yields. At some point, where data fails, it requires resorting to heroic assumptions. We will attempt to assess these assumptions as we go along and provide alternatives and caveats.

There have been some attempts to estimate these numbers including the well-known under-reporting of South Africa's agricultural sector in national accounts (Kirsten and Sihlobo 2021, Meyer 2022). Our aim here is to complement both the estimates of under-reporting and of potential crop agriculture output extending in some cases previous studies to the country as a whole so as to provide a sense of geographical distribution of the potential gains.

Our exercise is performed in two stages. First, we estimate under-reporting (production that is currently in place but not reported in the census or national accounts). Second, we estimate unmet productivity potential on currently utilized land.

Zero Hunger programme, 2011, the Ilima Letsema project, 2011, the 'one household one hectare' programme, 2015, the Agri parks initiative, 2015 and Operation Phakisa, 2017.

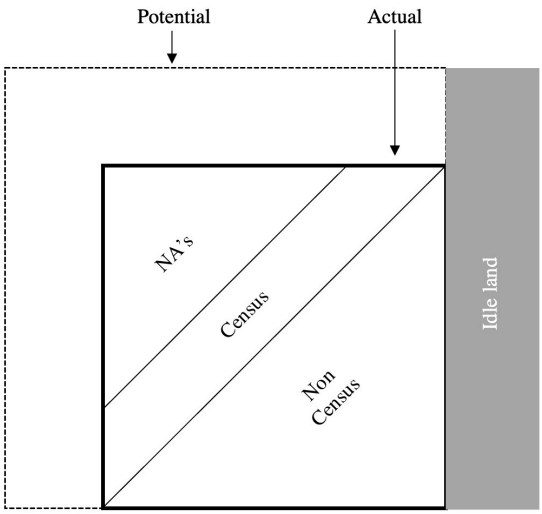


Figure 1: Overview

We can illustrate the scope of our exercise with the aid of Figure 1. In Figure 1, the solid line box can be understood as the total agricultural production. Part of this production is captured in South Africa's national statistics, the upper left triangle (NA's, for national accounts). But the census data shows that national statistics underestimate true production by a third (graph is not at scale), thus moving the observed production triangle in the box southeast. However, the census itself misses a large share of production. The remainder of the box is unaccounted agriculture both in the national accounts and the census, which needs to be estimated.

Once total actual production is obtained, we estimate how much larger the box may be based on estimates of potential productivity by crop. This extra crop production is indicated by the dashed lines that increase the size of the box. Our measure of productivity gap will be the difference between the larger square identified by the dashed lines and smaller solid line square box.

Notice that we are silent on the potential of unused land, land that today is not

being used for agriculture but could potentially be used. This is the issue of *idle land* and comprises abandoned land (including in the former homelands, as well as government land that remains inaccessible because of legal constraints and so forth). We will not provide an estimate of the potential of idle land in this exercise.

The computations provide stark conclusions. On the one hand we show that the under-reporting is larger than previously thought. If census data suggests that crop production is 1.4% of GDP we estimate it to be closer to 3.9% of GDP. However, and predictably, most of this underreporting is in horticulture. Field crops underreporting adds 0.75% of GDP while horticulture adds 1.7%. This is natural because horticulture production scale is smaller and therefore more likely to pass under the radar of the agricultural census.

Potential output is obviously larger. Using FAO's more conservative estimation crop production on currently utilized land could scale up to 4.9% of GDP. The jump relative to our estimate of actual production is evenly split between field crops and horticulture with both adding about half.

In short crop fields are underreported by .7% of GDP and .5% of GDP below potential. Horticulture is underreported by 1.7%, and below potential by 0.5% of GDP, though we will argue later this is our weakest estimation, and probably an underestimate.

We will show, in turn that this gains are very *unevenly* distributed across South Africa, suggesting the need for policies with a strong local component. From a policy perspective moving to potential could add about 350.000 jobs.

2 Under-reporting in Crop Agriculture

It has been argued that South Africa's agricultural sector is under-measured in national account statistics. Columns (1) and (2) in Table 1, replicated from Table 10.4 in Kirsten and Sihlobo 2021, shows that GDP data under-measures agricultural value added by around 30%, at least relative to census data (we will come back to the other columns of this table as we move along). While the national accounts suggest a participation of agriculture in the GDP of 2.3%, census data suggests this number should be increased to 3.4%. [3]

[3] All numbers in this table are in 2017 rands. Our base GDP is taken from Table 1.4 in Kirsten and Sihlobo 2021 at 4,243,794 million rands. The GDP numbers in the other columns are adjusted by the extra value added, except in column 7 where we do not adjust the GDP to evaluate the extra production relative to current output.

Table 1: Alternative measures of Agricultural Production

	Stats SA (1)	Census crop (2)	KS I (3)	KS II (4)	Census data (5)	BOS w/missing crop (6)	BOS potential crop (7)
Total Income	231,251	332,756	355,756	412,756	139,545	391,349	493,264
Intermediate Purchases	136,157	188,467	201,357	233,619	77,491	221,507	279,192
Net Value Added	95,094	144,289	154,399	179,137	59,419	169,846	214,072
% of GDP	2.3%	3.4%	3.6%	4.1%	1.4%	3.9%	4.9%
Adjusted GDP	4,134,521	4,243,794	4,253,904	4,278,642	4,243,794	4,353,075	4,353,075

In spite of this large gap, Kirsten and Sihlobo 2021 argue that even census data under-measures actual production. By comparing the number of farms in the census and in the 2011 Population Census and the 2016 community survey, they conclude that the 2017 Census may have missed about 214.800 farmers with a gross production income of between 23 and 80 billion rand. Adding these lower and upper estimates in columns (3) and (4) of Table 1, and keeping constant the ratio of income to intermediate inputs, would take the size of South Africa's agricultural sector to anywhere between 3.6% and 4.2% of GDP.

This under-reporting has its correlate in the labor market where subsistence farmers are not considered part of the labor force. CDE (2011)[4] estimates about 1.5 million unaccounted workers in the agricultural sector.[5] Shah 2022 using the 2019 Labor Market Dynamics Survey estimates this number closer to 1.9 million, that is, about 5% of the labor force. From every point of view this under-reporting is significant.

Can we provide an alternative estimation for the size of agricultural GDP? We can do so for the portion of agriculture that excludes livestock production, that is, crop production, including horticulture and perennial crops. We will call this group *crop farming* to match typical denomination.[6]

In order to estimate actual crop farming we follow a two-step procedure. We first measure the amount of land used for crop farming using satellite data. Once the amount of total crop farming land is obtained, we construct an estimate for actual yields on the portion that is unreported land. The combination of both a yield and a surface for unreported land will provide, when added to census data an estimate of total crop farm production. Note that by using land currently being used for production, this computation, as said, is silent about idle land, defined as land which

[4] Bernstein 2011

[5] See Table 1.

[6] See for example NAICS categories 1111 to 1119.

is not currently being used at all for agriculture though it potentially could be farmed.

2.1 Estimating total crop land

Our first step is to obtain a measure of the total amount of land used. In order to identify total crop land we use the *Global Food Security-Support Analysis Data Cropland Extent Product of Africa, 2017*[7] which is based on satellite data. This dataset uses satellite data from 2013 through 2016 and identifies lands cultivated with plants harvested for food, feed, and fiber, including both seasonal crops (e.g., wheat, rice, corn, soybeans, cotton) and continuous plantations (e.g., coffee, tea, rubber, cocoa, oil palms). It includes fallow lands, which are uncultivated during a season or a year but are farmlands and are equipped for cultivation.[8] [9]

A specific test of the accuracy of the satellite data can be obtained by comparison with reference data for specific areas. Table 2 shows a number of plots that are known to be used in crop agriculture and some which are known not to be. This reference data is then contrasted with what comes out of the satellite mapping. As can be seen, the mapping tends to slightly overestimate agricultural use as there are more units which are not being used for crop agriculture but are identified for this use, compared to those that are being used for crop production but are not identified for this use. The overall accuracy is about 95%, so we adjust our satellite estimates by this amount to account for the overestimate.

The amount of crop land measured by satellite data is significantly higher than that obtained from the census. This is natural as census data only obtains informa-

[7] See Oliphant, Thenkabail, and Teluguntla 2022; https://lpdaac.usgs.gov/products/gfsad30afcev001/.

[8] The reason the computation uses a three year interval is so as not to discard fields that on a specific year may not be used because of rotation. Only if the land is not used in any of the three years it is deemed non agriculture land.

[9] All the computations in the paper can be replicated with a python code that can be downloaded from https://drive.google.com/file/d/1kd2iIvtqzuhmWYy09Dyi8t2-Wf14cwRW/view?usp=share$_l ink$.

Reference Data			
		Crop	Non-Crop
Map data	Crop	22	9
	Non-Crop	4	215

Source: GFSAD

Table 2: Accuracy of satellite data

tion from firms that are registered as agricultural producers in the VAT database and choose to respond to the survey. Therefore, it not only misses those producers that do not respond the survey but also misses out on smaller units that may be registered under another activity, as well as informal producers. In our case the difference between the two across South Africa as a whole adds up to 10 million hectares which is sizable considering that census area is 7.6 million hectares.[10] Of these 10 million hectares 2.4 are located in former homelands.

Figures 2 and 3 show how the satellite data works in mapping agricultural land. The data is from the Kagisano/Molopo municipality in Dr Ruth Segomotsi Mompati district in North west province. We've chosen to show a picture from this area because we find a massive discrepancy between census and satellite data in this region. The pictures allow to see clearly that there is extensive agriculture in the area.

Figure 4, depicts visually where crop land is found in South Africa. The map is very intuitive and well known and shows that the activity is located mostly in the eastern part of the country as well as in the Cape region with the arid northwest mostly lacking crop production (except in contained river valleys). Table 3 shows how this surface is distributed across regions.

This discrepancy is not new to South African agricultural economists. Not only

[10]See Table G in https://www.statssa.gov.za/publications/Report-11-02-01/Report-11-02-012017.pdf.

Figure 2: Satellite image

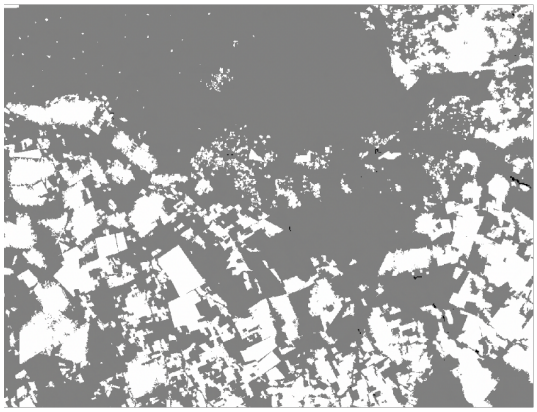


Figure 3: Satellite Mapping

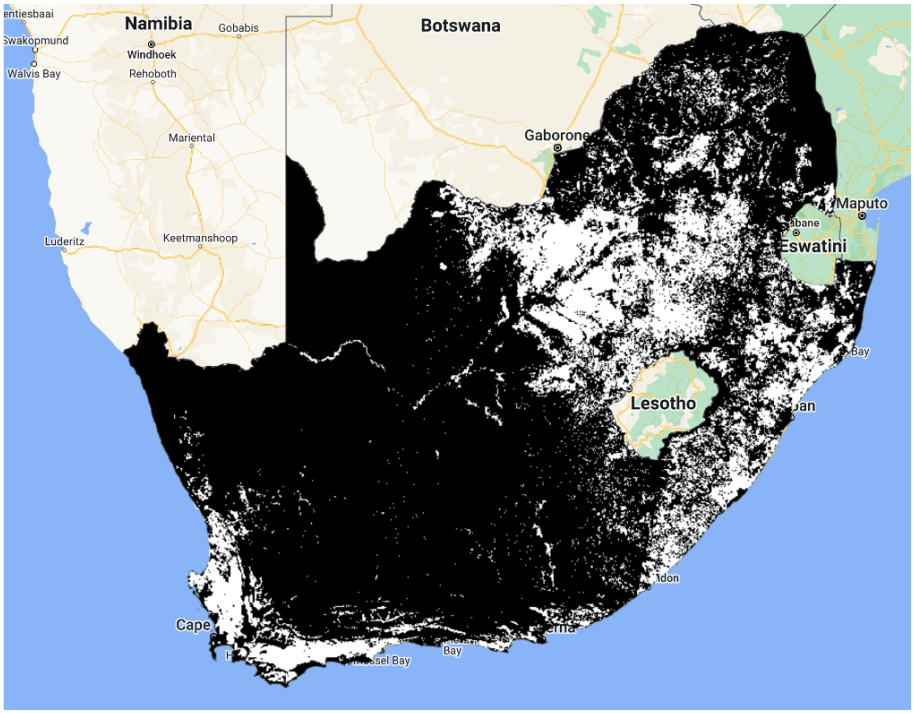


Figure 4: Satellite Crop production surface

	Census Arable Land Hectares	Additional Identified Hectares
Eastern Cape	357,809	1,386,081
Free State	2,454,123	1,499,147
Gauteng	180,349	295,872
KwaZulu-Natal	524,053	1,830,197
Limpopo	361,341	895,516
Mpumalanga	943,163	551,570
North West	1,118,508	2,016,075
Northern Cape	671,400	18,993
Western Cape	1,003,643	1,398,615
Total	7,614,389	9,892,066

Table 3: Census and Satellite crop surfaces

was it already mentioned in Kirsten and Sihlobo 2021, but previous aerial surveys have also identified such discrepancies. For example those in the Limpopo province had identified massive under-reporting in the census data. Table 4, provided to us kindly by Ferdi Meyer, makes this point clearly (Meyer 2022).[11]

While there was never an o cial report by the Provincial Department of Agriculture in Limpopo of this data, they did share some of the findings from the fly-over census with Meyer's team who put together Table 4. Table 4 allows a comparison of these aerial surface estimates, with the o cial census data from 2007, at the time the most recent census data available.

Notice that in Table 4 the total hectares registered by the aerial census is about double the actual Stats SA data (119,395 vs 58,426). This ratio is smaller than the one we find for Limpopo itself from our satellite data (1,256,000 vs 361,000) in recent years, but not that different from the one we obtain for the whole of South Africa (17,506 vs 7,614), where we find that while census data identified 7.6 million hectares of arable land, actual land under cultivation adds 10 million hectares.

[11]Estimates from Subtrop.

Crop	Aerial Census 2012 (Ha)	Stats SA 2007 (Ha)	% Diff	Underestimation (million rands)
Wheat	32,218	12,985	148	78
Potato	14,197	8,526	66	743
Tomato	12,243	4,711	159	3,163
Cabbage	3,034	237	1180	572
Avocado	10,811	7,568*	42	2,432
Mango	10,813	4,756*	127	1,514
Citrus	36,079	19,643	83	2,350
Total	119,395	58,426	104	10,855

Source: GFSAD

Table 4: Limpopo's aerial survey

Figure 5 shows the amount of total satellite crop land relative to census arable land. Notice that the former homelands tend to show the highest ratio, and that most of the under-reporting comes from the eastern frontier. However, unreported land is spread out across the country, except in the northwest where, because of its arid climate, has little crop potential and where the firms are probably formal and more likely to answer the census.

2.2 Estimating the yields on unaccounted land

Once we obtain an estimate for the the total amount of cropland we need to assign a yield to the unreported surface. One option would be to just extend the average yield in the census to the non-census surface. But this computation would miss the fact that census land corresponds, most likely, to the more productive areas. This is due to historical legacies and endogeneity: well established firms, more likely to respond the census, are probably located in the most convenient land. So extending the census yields to non reported land could lead to a very large overestimation of the value of unreported agricultural production.

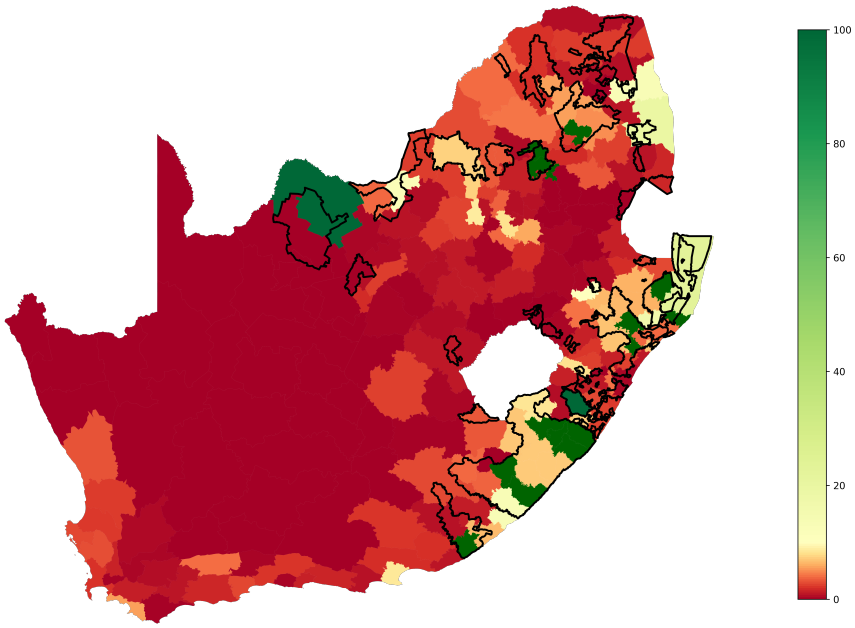


Figure 5: Ratio of crop land to census land

In order to solve this problem we utilize the FAO's potential yield estimation[12]. FAO provides a measure of potential yields for a wide array of crops with a granularity of roughly 81 square kilometers covering the whole of South Africa's territory.

FAO has potential dry weight yields for staple crops, nuts and seeds[13]. FAO provides a range of productivity alternatives and we have taken those corresponding to natural rainfall which we think is more appropriate for a country-wide analysis.

After converting dry weight to fresh weight we can express crop weight in dollar value using FAO 2000 international crop prices providing an estimate of potential

[12]Fischer et al. 2021

[13]More specifically, banana, barley, sorghum, groundnuts, maize, potatoes, oats, oil seeds, sugarcane, sunflower, soyabeans, sweet potatoes and wheat. Unfortunately it does not include most fruits and vegetables including grapes, which lowers the quality of our estimate in the Western Cape.

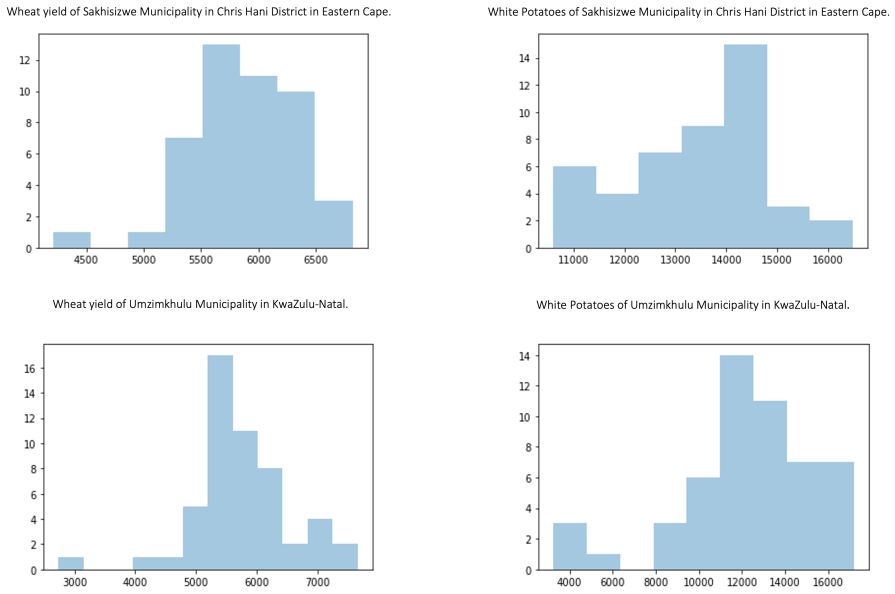


Figure 6: Yield Distributions Within District

yield in USD/ha for each crop. Given that the unit size of FAO's computation is small, we have many FAO's potential yield values in dollars per hectare within each municipality. In fact, we can compute a whole distribution of yields across a municipality. As an example, Figure 6 shows histograms of yields for wheat and white potatoes for the Sakhisizwe and Umzimkhulu municipalities.

Once we have the distribution of yields within each municipality, we make the natural assumption that census land has been allocated to the best land. So, for example, if in a given municipality census land is equivalent to 50% of total estimated land, we assume it is the 50% best land, and it's yield corresponding to the 50% upper range of yields. We then take for each crop the average yield of the highest 50% and compare to the average yield of the lowest 50%. Once we have the ratios for each crop we weight these averages by the actual weights of each crop within the municipality. If we don't have information for a specific crop, we drop

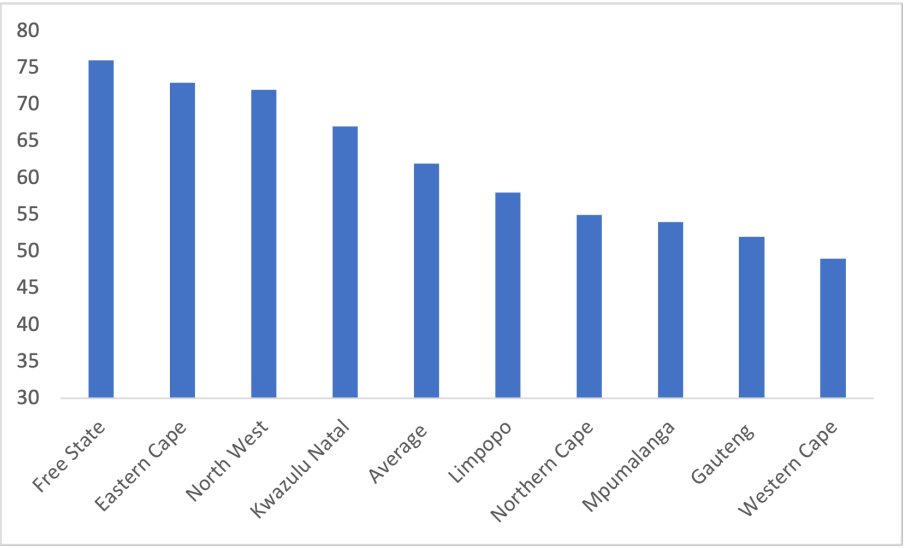


Figure 7: Non-census vs census yields across provinces

it and re-weight the shares with crops for which we have data.[14] Figure 7 shows the distribution of these relative yields across regions. The estimates indicate that, using this criteria, potential yields in the unaccounted land range from 75% of census land in the Free State, to a low of around 50% in the Western Cape. The yields for unaccounted land are higher if the province is geographically more homogeneous. With sharp differences in climate and land properties it makes sense to find larger differences. In the two ends of the chart we find the Western Cape that has the most diverse rainfall patterns within any province, and the Free State that has a more homogeneous rainfall, giving some intuitive support to the estimates.

We use the actual 2017 Census data to compute a weighted average (based on crop surface) yield in each municipality. With this average yield in each municipality and the relative yield for non census land, it just remains to apply this ratio to the yield estimate of census land to obtain an estimate of the potential value per hectare

[14] An alternative would be to average the yields in each FAO square using municipality weights and then look at the distribution of these yields. Both methods should give similar results if there is not excessive segmentation of crops within each municipality.

for non-census land. More concisely, the value per hectare of unreported land is $(VPHUL)$:[15]

$$VPHUL = \frac{\text{Census Agriculture Income}}{\text{Census Agriculture Land}} \frac{\text{Non census Land Yield}}{\text{Census Land Yield}}. \quad (1)$$

The first ratio provides the census estimate of the value of agricultural production per hectare, which is income divided by surface. The second ratio is the relative yield of non-census to census land estimated as explained above.

In short, the value of production of non-census land, which we call missing value of crop production $(MVCP)$:

$$MVCP = \sum_m (\text{Total Crop Land}_m \quad \text{Census Land}_m) * VPHUL_m, \quad (2)$$

where m aggregates across municipalities.

In homelands, for which there is no census agriculture reported and therefore no distribution of yields, we use the minimum average yield from all of the neighboring regions.

2.2.1 Splitting the Missing value

We just reviewed a way of estimating an aggregate missing value of crop production. But this missing value corresponds to two big universe of products: field crops and horticulture. Field crops typically involve more extensive production, little irrigation if any, whereas horticulture is typically concentrated in smaller surfaces and are supported by more infrastructure, and therefore less subject to climatic conditions.

[15] In some cases, due to border overlaps, we need to redistribute some satellite lands from neighboring municipalities with excess satellite land to municipalities that need more satellite land to get at least census arable land amount. For some municipalities, satellite cropland underestimated cropland, in that case we just used actual cropland estimation.

Our relative yield estimate is mostly based on extensive field crops both for perennial and non perennial crops, so has little if anything to say about the relative productivity of unreported land in horticulture. While it is reasonable to assume that field crops in unreported land is produced in smaller units or less suitable regions and productivity smaller. For horticulture, the reason it goes undetected is not productivity but its size so that the argument that there is lower productivity in land missed by the census is not as clear. On the other hand, if undetected horticulture corresponds to informal farmers and detected land to large agricultural companies, there is a way of arguing that undetected horticulture production may have a lower productivity.

In the following exercise we have assumed that the relative productivity of horticulture in the surface not captured in the census, has the same relative productivity than that of field crops. This is a working assumption. If the reader thinks there is no real reason for a change in productivity, our numbers for horticulture should be considered an underestimate. If the reader thinks there is a larger productivity difference in formal and informal production of horticulture than in crop fields, our estimates will be an overestimate.

In order to compute the missing value of unreported land we assign to it the same ratio of horticulture and field crops that we have in each province on census land. This assumption may also not be correct if most of the unreported surface is in horticulture. If so this assumption would imply that our estimates are an underestimate.

With the horticulture surface and the common relative yield we can compute the missing value in the two categories. Beyond the readers assessment of the value of horticulture (which really requires heroic assumptions) it is important to split out horticulture, in order to assess the missing value in field crops for which our estimates are more precise.

2.3 Estimating missing crop farming value

Table 5 provides our estimates of missing crop farming value. The left box in panel (a) provides census total income by province from the Census Table 1.4 which is our starting point. The center box in the same panel provides our estimate of missing crop income by province. Whereas crop income adds up to 140 billion rand, missing value adds 252 billion. Notice however, that most of this missing value is in horticulture (172 billion) whereas crop fields underestimation is much smaller at 75 billion (which still implies duplicating census estimates). The geographical distribution is also uneven. While there is little underestimation in the Northern Cape, it is massive in Kwazulu Natal or in the Northwest. In both these regions the census missed out completely on horticulture. [16]

Panel (b) transforms the numbers of panel a into value added, using Kirsten and Sihlobo's ratio of income to value added. While census data suggest a contribution to GDP of 60 billion rand, we add around 109 billion to a total of 169 billion. As in panel (a) the biggest effect is in horticulture. Missing crop value in field crops is 32 billion but missing value in horticulture is 74 billion.

Panel (c) measures these numbers in percentages of GDP. While census value added in agriculture is estimated at 1,4% of GDP we propose a 3,9% measure. Field crop contribution to GDP is 0.7% in o cial numbers, we estimate it to be 1.4%. Horticulture contribution is 0.7% we estimate it to be 2.4%.

Table 5 shows where this missing agricultural income is bigger. The North West appears as a big outlier, so large that we believe this speaks more about the quality of the census in this region. More predictably KwaZulu-Natal and Limpopo appear as regions with large undereporting. The two regions with the lowest underreporting

[16] The table presents the numbers by province. In order to obtain a GDP per province for our adjusted GDP number we just applied the ratio of each province GDP in national accounts. This implies that our 4,353,075 2017 million rand GDP is split in 335,469 for the Eastern Cape, 219,362 for the Free State, 1,490,950 for Gauteng, 698,164 for Kwazulu-Natal, 318,300 for Limpopo, 326,451 for Mpumalanga, 282,009 for the North West, 90,256 in the Northern Cape and 592,115 for the Western Cape.

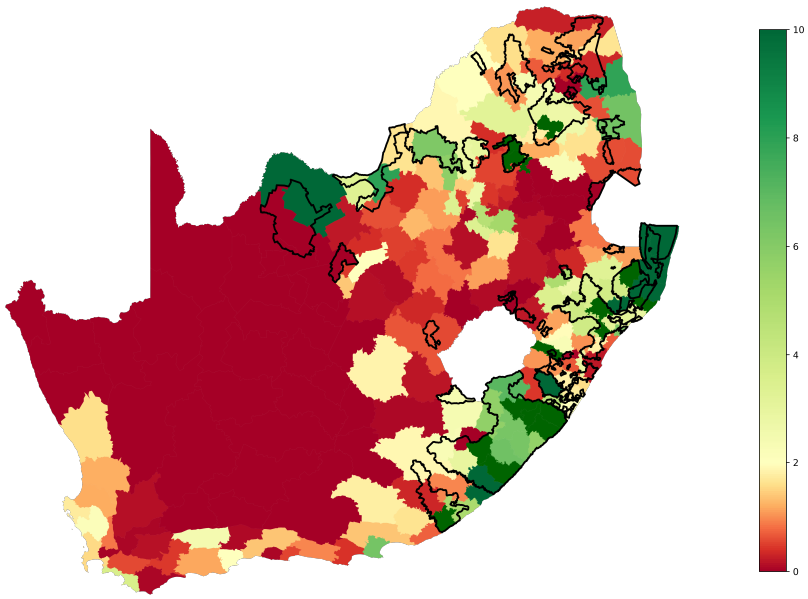


Figure 8: Ratio of missing and current crop income

are Gauteng and the Northern Cape, with low informality the first and due to low crop production the second.

In Table 1 columns 5 and 6 put these estimates in perspective with existing estimates. Because we are dealing with crop land we re-estimate the census numbers only considering the items in farming products.[17] In column 5 this provides an estimate of 1.4% of GDP participation of crop farming production (see also panel c in Table 5). Our estimate adds 252 billion to gross income (see panel a of Table 5). Keeping the ratio of intermediate goods constant it delivers an estimate of current agriculture production of 3.9% of GDP.

Figure 8 shows the geographical distribution of unreported crop production across South Africa as a share of GDP. With the exception of the northwest which is arid

[17]See Table 1.4 in the agricultural census.

and not suitable for production, the loss is quite widely distributed across the country, with the exceptions of some pockets where commercial agriculture is dominant. At any rate, it is clear that the census has under-reported more significantly in the east and the north.

The difference between actual production and census production adds up to 252 billion rand in gross income and 109 billion in value added. Current employment for the corresponding categories of the census is 502 thousand workers. Scaling this up to our estimates brings the number to 1,4 million workers. These numbers are quite similar to those mentioned in the introduction so they provide an independent validation for our MVCP estimate.

2.3.1 How does this compare with Kirsten and Sihlobo's missing value estimate?

As we mentioned above Kirsten and Sihlobo 2021 estimate a missing value of agricultural production of 23 and 80 billion. Given that our estimates are larger than theirs, we compare ours with their 80 billion upper bound.

Using Kirsten and Sihlobo's decomposition of agricultural production in livestock, field crops and horticulture presented presented in their Table 10.2 (which splits agriculture 48% in livestock, 24% in crops and 28% in horticulture), their estimate of 80 billion of gross income for agriculture translates into an under-reporting of 19.2 billion in field crops and of 22.4 billion in horticulture. Our equivalent numbers are 75 billion and 171 billion, substantially higher.

Why are these numbers so different? Kirsten and Sihlobo estimate is based on the number of potential firms missed by the census. So they do not pretend to capture subsistence farming, not even the complete set of farms. Our estimates, on the contrary, are driven not by the number of farmers, but from the satellite identified land

which, as we mentioned, more than doubles census land. It is not surprising then that our number more than doubles census gross income in line with the increase in reported land. In doing so we are including all types of farmers even subsistence production which probably increases significantly the number.

While we feel confident that an estimated missing production should consider all land used in crop production, we need to be aware that we are also attaching to these other farmers (including subsistence farmers) a relative productivity that corresponds to what FAO considers may be different productivity according to geographical and climatic differences, i.e. without considering *technological differences in production*. This implies that we are assuming subsistence farmers produce as "e ciently" as commercial producers, even though they produce less because they do so in a poorer land.

Covering the whole land under cultivation is probably an improvement in the attempt to measure missing agricultural production, but adjusting the productivity of land not reported in the census only for climatic and land quality differences ignoring productivity differences, may lead to an overestimation, particularly in field crops.

Probably the true number lays somewhere in between both estimates. At any rate, by splitting the numbers, each research can adjust the missing value by their preferred "productivity gap" coe cient (we have chosen not to do this here) to obtain a new estimate, an exercise that is easy to do and may even be different by region.[18]

3 Estimating South Africa's productivity gap

Our next computation will provide an estimate of how far South Africa's crop production is from it's frontier. In terms of Figure 1, we want to estimate the area comprised between the dashed lines and the square box. How bigger could South

[18] If such an adjustment is made, it would reduce missing agriculture but increase the gap between actual and potential which will remain unchanged.

Africa's crop production be?

For this computation we will use FAO's potential yield estimates and apply them to *the entire* crop production surface. This number should be taken with care. On the one hand we have taken FAO's productivity with natural rainfall which provides a conservative approach by not considering the potential that could be obtained with irrigation. However FAO's potential assumes e cient use of inputs such as fertilizers, relatively e cient infrastructure, availability of financing etc. In this sense, the number is not a measure of South Africa's current possibilities, but rather of its physical potential. We believe it is still a useful benchmark as it indicates what greater policy focus could yield.

In order to estimate this potential, we already have most elements. Using FAO's potential yield per crop we can compute an average yield for each crop in each municipality (though only for the crops for which we have a potential yield estimate). We then calculate a weighted average yield as a weighted average (per crop surface) for all available crops within the municipality. We then compare this average yield with the average yield from census data, also for the same crops. The ratio of potential to actual will give a number larger than one. By applying this to the agricultural production in the municipality we obtain a measure of potential agriculture.

Of course, applying one common "productivity gap" coe cient to all production is questionable. For example, in the Western Cape, this yield differential is on crops but does not include grapes. So applying the yield difference means applying it to grape production, though there is no strong reason to argue that the productivity gap in other crops should extend to such a specific crop as grape production. An alternative, but equivalent arbitrary computation would be to leave these crops for which we have no information out of the computation. A similar caveat applies to horticulture, where extending the productivity gap of extensive crops also finds weak support. This is an additional reason to split split the computation for field crops and for horticulture, so the reader can disregard the computation for horticulture if

it feels fit to do so.

Aware of these caveats we plow ahead. The results are presented in Table 6. Our estimate for potential crop production is 493 billion rand, of which 252 billion corresponds to the former homelands meaning that there is a 100 billion (in 2017 rands) difference between potential and our measure of actual agriculture. This difference is about 1% of GDP.

In contrast to our measurement of missing agricultural value, the slack relative to potential is more evenly spread between field crops and horticulture (for which we argue our number has less support). The distance from the frontier for field crops is equivalent to around 0.5% of GDP.

Table 6 shows these numbers by region. Focusing on field crops exclusively we find that there is a significant geographical dispersion in the distance to potential. 2.7% of GDP in the Free State, 1.1% in Kwazulu Natal, 0.7% in the Western Cape and 0.6% in Mpumalanga. In the other regions the distance is 0.5% or lower.

Figure 9 compares actual production and distance to the frontier. The horizontal axis shows the current contribution of field crop production to GDP and the vertical axis the distance to the frontier. Agricultural policies should take into consideration the relevance of the agricultural policy and the potential effectiveness of such policies.

The Free State appears to the right because there the contribution of field crops to GDP doubles that of any other province in the country. The potential extra GDP is also, naturally, larger. However what can be added is substantially below actual, as the region is a significant hub for commercial agriculture.

The Eastern Cape is quite different. It shows a smaller contribution of crop field to GDP, but its potential is almost equivalent to current production (this is the reason the dot for the Eastern Cape appears on the 45^o line).

The graph suggests the Eastern and Western Capes, Kwazulu-Natal and the Free State as the places where more potential is to be added. On the other end Gauteng, Northern Cape, the Nortwest Province and (suprisingly) Limpopo appear as close to their natural potential.

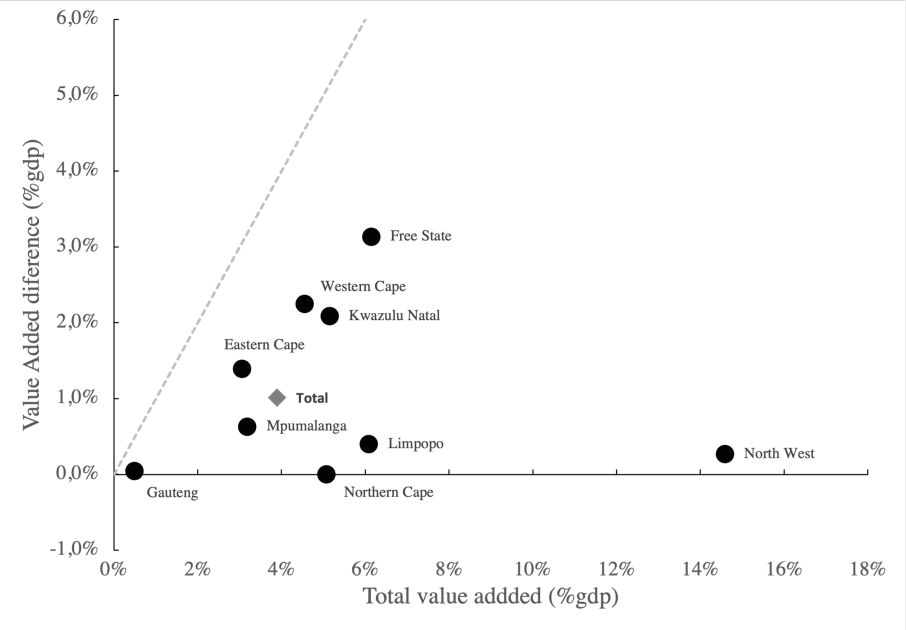


Figure 9: Actual and Potential Contribution to GDP

To put this into the perspective of the other computations discussed so far, we include these computations in the final column of Table 1. Maintaining the ratio of intermediate inputs this provides a value added for crop production of 214 billion rand, 4,9% of GDP, with a difference of 100 billion 2017 rands of unexploited potential when measured in terms of gross income. Again, applying a proportional output-employment ratio this difference would imply close to 350.000 additional jobs.

Figure 10 shows the distribution of potential output, as the crop income is divided by municipality surface, lighter and greener areas indicated those in which

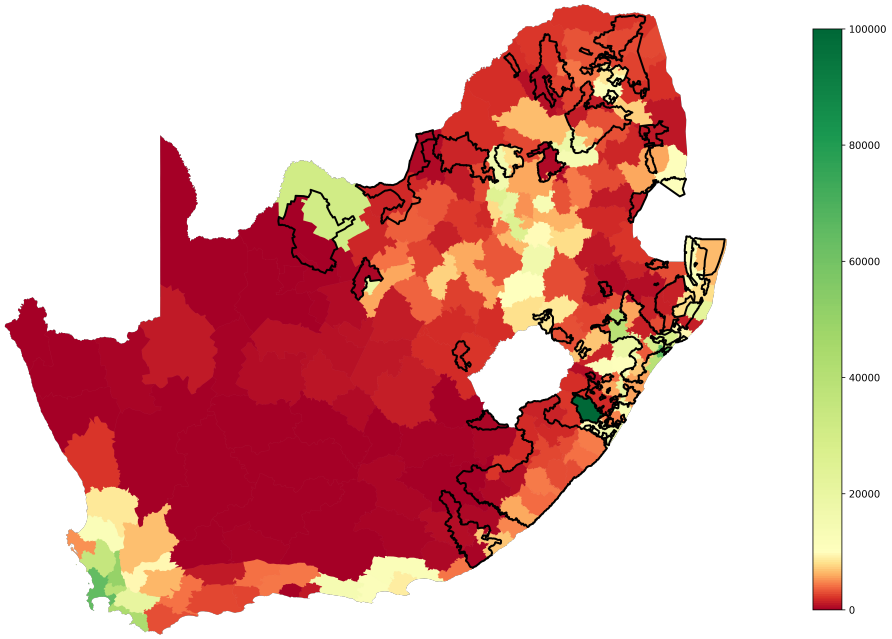


Figure 10: South Africa's Potential Yield over entire municipality surface

agriculture is more relevant at the municipality level.

Figure 11 shows the geographical distribution of potential production relative to our estimates of actual production by municipality. Surprisingly the biggest differences are in a corridor running for Johannesburg towards the south including the Kwa-Zulu Natal province. The north west, the north and the eastern rim along the Mozambique border appear close to it's potential.

4 Conclusion

This paper has produced an estimate of undereporting of South Africa's crop production and an estimate of how far away South Africa's crop production is from its frontier. The computation requires, at points, resorting to heroic assumptions,

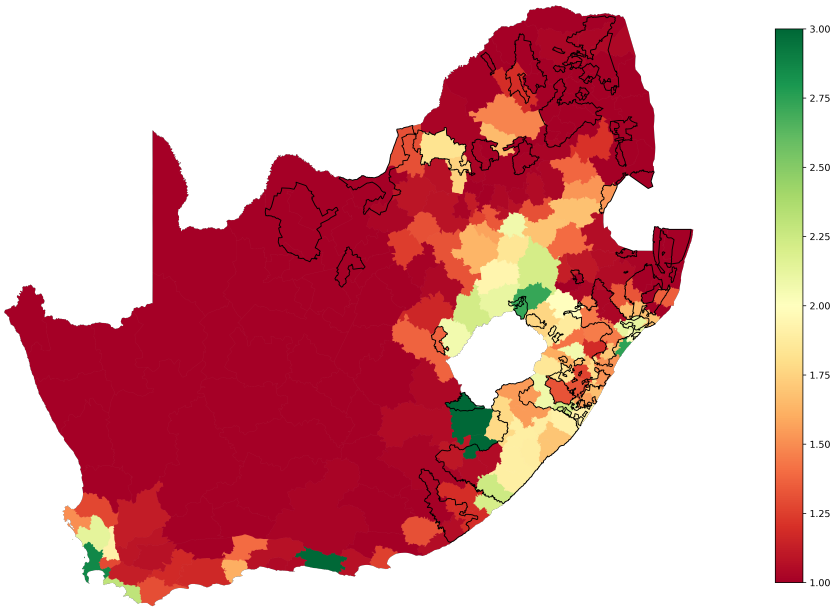


Figure 11: Ratio of potential over missing and current crop income

but has the virtue of building on a satellite record of all land used for crop production.

The first result suggests that South African national accounts can benefit from available satellite data to produce better estimates of this underreporting that, at 2,5% of GDP we find is larger than previously thought. Most of this underreporting appears naturally from horticulture which is more di cult to track than extensive corps. In fact 1.7% of that 2.5% is explained by horticulture.

Second, we have detailed that despite South Africa's clear capabilities in the agricultural sector, there remains unexploited potential after adjusting for under-reporting in national accounts and in the census. This potential is close to 100 billion rand of gross income which translates to somewhat less than 1% percent of GDP. Half of which comes from field crop production and the rest from horticulture. Adding this extra 1% of GDP would allow adding about 350.000 new jobs.

In our estimation we have assumed that non reported land produces with the same technology as commercial land, except that adjusted by land and climatic conditions of that land. This leads to large number of undermeasurement and a smaller gap between current production and potential. If we choose to adjust the productivity of unreported land to adjust for productivity changes, this will reduce the size of unreported crop production but increase the gap with potential. For example for field crop if we adjust the productivity of unreported land to half (say, because subsistence farmers don't have access to the machinery or inputs that commercial agriculture uses), this would bring down the missing field crop production from 0.7% of GDP to 0.35% of GDP. At the same time it will increase the gap with potential from 0.5% of GDP to 0.85% of GDP. Our estimates are presented in su ciently granular form that this adjustment can be done according to the preferences of the researcher. They can also be estimated easily using the python code provided above.

This potential, however, is not evenly distributed across the territory. Knowing where this untapped potential is allows to focus and target crop development policies from a policy perspective. We find that a significant part of South Africa's missed potential are in regions where commercial agriculture is strong. This suggests that agricultural policies should not miss out a discussion of constraints to commercial agriculture and particularly of the weakening of property rights there. Former homelands also suggest untapped potential, but the issue does not, according to our estimates, apply equally to all homelands. Homelands in the north of the country are, according to our estimates, at potential, so attempts to bring up production there are prone to failure. Kwazulu-Natal and the South Eastern homelands can provide a better return for the policies implemented there.

In short, agricultural policies can not be implemented as a blanket approach. Beyond the debates on land re-distribution and traditional property rights reforms there is a rich policy agenda covered in detail by the Agricultural Master Plan (NAMC 2022). Yet South Africa has a decades-long history of detailed master plans leading to little or no results. Capacity constraints in the public sector persist, fiscal

resources are under increased strain, and broader political debates regarding land ownership are unresolved.

So learning to target the efforts is key. We hope this paper allows to at least bring awareness to the issue of the need for continued policy focus and geographical differentiation of policies. In this endeavor it is also important for other actors, in particular the private sector and civil society, to explore what they can do in the meantime to expand agricultural output toward its potential.

References

Beinart, William and Peter Delius (2018). "Smallholders and land reform: A realistic perspective". In.

Bernstein, Ann (2011). "A fresh look at unemployment: A conversation among experts". In.

Cousins, B (2018). "Employment in informal-sector agriculture in South Africa". In: *The South African Informal Sector: Creating jobs, reducing poverty. Cape Town: World Rights*.

Fischer, Gunther et al. (2021). *Global Agro-Ecological Zones v4–Model documentation*. Food & Agriculture Org.

Kirsten, J and W Sihlobo (2021). *The Oxford Handbook of the South African Economy*. Oxford University Press. Chap. Agriculture in South Africa.

Meyer, F (2022). "Table based on report by the Provincial Department of Agriculture, Limpopo and StatsSA". In: *Presentation, Bureau for Food and Agricultural Policy*. Mimeo.

NAMC, BFAP CCRED (2022). *Agriculture and Agro-Processing Master Plan*. National Agricultural Marketing Council, Bureau for Food, Agricultural Policy Center for Competition, Regulation, and Economic Development. Published by Ministry of Agriculture, Land Reform Rural Development, Republic of South Africa.

Oliphant, Adam, Prasad Thenkabail, and Pardhasaradhi Teluguntla (2022). *Global food-security-support-analysis data at 30-m resolution (GFSAD30) cropland-extent products—Download Analysis*. Tech. rep. US Geological Survey.

Shah, K (2022). "Labour Markets and South Africa's Triple Challenge". In: Mimeo.

(a) Gross Income

	Census			Missing			Ratio of Missing to Census		
	Fields	Horticulture	Total Crop	Fields	Horticulture	Total Crop	Fields	Horticulture	Total Crop
Eastern Cape	1,744,725	6,712,227	8,456,952	2,775,587	9,903,309	15,247,880	1.6	1.5	1.8
Free State	17,892,194	2,423,161	20,315,355	9,896,377	836,237	10,826,970	0.6	0.3	0.5
Gauteng	3,790,882	3,684,161	7,475,043	3,236,875	6,059,296	9,426,262	0.9	1.6	1.3
Kwazulu Natal	9,097,909	2,598,547	11,696,456	27,916,120	42,493,750	71,319,680	3.1	16.4	6.1
Limpopo	4,452,743	16,265,755	20,718,498	6,892,489	15,931,120	24,037,560	1.5	1.0	1.2
Mpumalanga	10,794,723	5,158,220	15,952,943	4,560,326	3,684,641	8,062,617	0.4	0.7	0.5
North West	7,419,599	2,514,860	9,934,459	10,835,600	74,237,110	85,084,220	1.5	29.5	8.6
Northern Cape	4,661,189	4,892,900	9,554,089	656,034	366,073	1,022,106	0.1	0.1	0.1
Western Cape	9,175,479	26,284,389	35,459,868	8,367,585	18,391,460	26,759,050	0.9	0.7	0.8
Total	69,029,443	70,534,220	139,563,663	75,136,993	171,902,996	251,786,345	1.1	2.4	1.8

(b) Value Added

	Census Value Added			Missing Value Added			Total Value Added		
	Fields	Horticulture	Total	Fields	Horticulture	Total	Fields	Horticulture	Total
Eastern Cape	755,466	2,906,394	3,661,860	1,201,829	4,288,133	6,602,332	1,957,295	7,194,527	10,264,192
Free State	7,747,320	1,049,229	8,796,549	4,285,131	362,091	4,688,078	12,032,451	1,411,320	13,484,627
Gauteng	1,641,452	1,595,242	3,236,694	1,401,567	2,623,675	4,081,571	3,043,019	4,218,917	7,318,265
Kwazulu Natal	3,939,395	1,125,171	5,064,565	12,087,680	18,399,794	30,881,421	16,027,075	19,524,965	35,945,986
Limpopo	1,928,038	7,043,072	8,971,110	2,984,448	6,898,175	10,408,263	4,912,486	13,941,247	19,379,373
Mpumalanga	4,674,115	2,233,509	6,907,624	1,974,621	1,595,450	3,491,113	6,648,736	3,828,959	10,398,737
North West	3,212,686	1,088,934	4,301,621	4,691,815	32,144,669	36,841,467	7,904,501	33,233,603	41,143,088
Northern Cape	2,018,295	2,118,626	4,136,921	284,063	158,509	442,572	2,302,358	2,277,135	4,579,493
Western Cape	3,972,982	11,381,140	15,354,123	3,623,164	7,963,502	11,586,669	7,596,146	19,344,642	26,940,792
Total	29,889,749	30,541,317	60,431,066	32,534,318	74,433,997	109,023,487	62,424,067	104,975,314	169,454,553

(c) As share of GDP

	Census Value Added			Missing Value Added			Total Value Added		
	Fields	Horticulture	Total	Fields	Horticulture	Total	Fields	Horticulture	Total
Eastern Cape	0.2%	0.9%	1.1%	0.4%	1.3%	2.0%	0.6%	2.1%	3.1%
Free State	3.5%	0.5%	4.0%	2.0%	0.2%	2.1%	5.5%	0.6%	6.1%
Gauteng	0.1%	0.1%	0.2%	0.1%	0.2%	0.3%	0.2%	0.3%	0.5%
Kwazulu Natal	0.6%	0.2%	0.7%	1.7%	2.6%	44%	2.3%	2.8%	51%
Limpopo	0.6%	2.2%	2.8%	0.9%	2.2%	3.3%	1.5%	4.4%	6.1%
Mpumalanga	1.4%	0.7%	2.1%	0.6%	0.5%	1.1%	2.0%	1.2%	3.2%
North West	1.1%	0.4%	1.5%	1.7%	11.4%	13.1%	2.8%	11.8%	14.6%
Northern Cape	2.2%	2.3%	4.6%	0.3%	0.2%	0.5%	2.6%	2.5%	5.1%
Western Cape	0.7%	1.9%	2.6%	0.6%	1.3%	2.0%	1.3%	3.3%	4.5%
Total	0.7%	0.7%	1.4%	0.7%	1.7%	2.5%	1.4%	2.4%	3.9%

Table 5: MVCP per province

(a) Actual and Potential Gross Income

	Crop income			Potencial income			Difference		
	Fields	Horticulture	Total Crop	Fields	Horticulture	Total Crop	Fields	Horticulture	Total Crop
Eastern Cape	4,520,312	16,615,536	23,704,832	8,145,247	25,135,260	34,476,370	3,624,935	8,519,724	10,771,538
Free State	27,788,571	3,259,398	31,142,325	41,466,030	5,363,225	47,020,610	13,677,459	2,103,827	15,878,285
Gauteng	7,027,757	9,743,457	16,901,305	7,822,021	10,357,340	18,350,750	794,264	613,883	1,449,445
Kwazulu Natal	37,014,029	45,092,297	83,016,136	55,063,410	61,062,990	116,668,400	18,049,381	15,970,693	33,652,264
Limpopo	11,345,232	32,196,875	44,756,058	12,869,850	32,670,020	47,681,340	1,524,618	473,145	2,925,282
Mpumalanga	15,355,049	8,842,861	24,015,560	19,901,990	9,195,612	28,760,780	4,546,941	352,751	4,745,220
North West	18,255,199	76,751,970	95,018,679	19,485,580	76,336,530	96,761,970	1,230,381	-415,440	1,743,291
Northern Cape	5,317,223	5,258,973	10,576,195	5,317,203	5,258,974	10,576,180	-20	2	-15
Western Cape	17,543,064	44,675,849	62,218,918	27,585,760	65,383,190	92,968,950	10,042,696	20,707,341	30,750,032
Total	144,166,436	242,437,216	391,350,008	197,657,091	290,763,141	493,265,350	53,490,655	48,325,925	101,915,342

(b) Gap to Frontier as Share of GDP

	Percentage Change			Vale added of Difference			Percentage of GDP		
	Fields	Horticulture	Total Crop	Fields	Horticulture	Total Crop	Fields	Horticulture	Total Crop
Eastern Cape	80.2%	51.3%	45.4%	1,569,597	3,689,040	4,664,076	0.5%	1.1%	1.4%
Free State	49.2%	64.5%	51.0%	5,922,340	910,957	6,875,297	2.7%	0.4%	3.1%
Gauteng	11.3%	6.3%	8.6%	343,916	265,811	627,610	0.0%	0.0%	0.0%
Kwazulu Natal	48.8%	35.4%	40.5%	7,815,382	6,915,310	14,571,430	1.1%	1.0%	2.1%
Limpopo	13.4%	1.5%	6.5%	660,160	204,872	1,266,647	0.2%	0.1%	0.4%
Mpumalanga	29.6%	4.0%	19.8%	1,968,825	152,741	2,054,680	0.6%	0.0%	0.6%
North West	6.7%	-0.5%	1.8%	532,755	-179,886	754,845	0.2%	-0.1%	0.3%
Northern Cape	0.0%	0.0%	0.0%	-9	1	-6	0.0%	0.0%	0.0%
Western Cape	57.2%	46.4%	49.4%	4,348,487	8,966,279	13,314,764	0.7%	1.5%	2.2%
Total	37.1%	19.9%	26.0%	23,161,454	20,925,126	44,129,343	0.5%	0.5%	1.0%

Table 6: Census income and potential income income